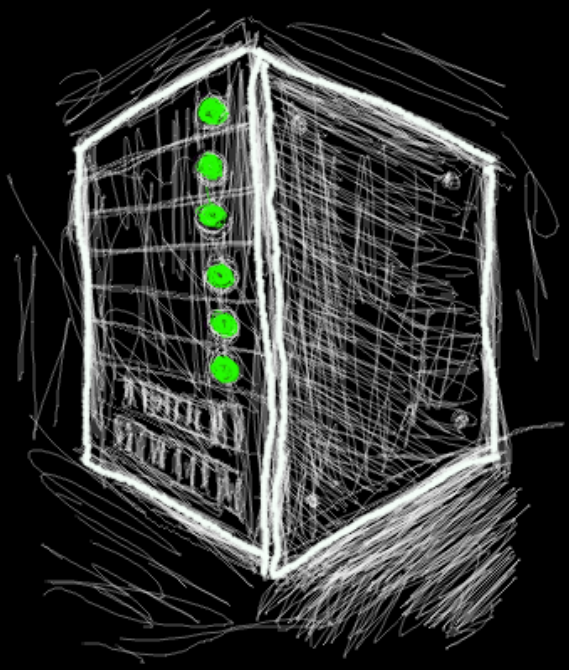
LET'S FORGET ABOUT IT

The Web of problems for the right to be forgotten

Paulan Korenhof



Let's forget about it. The Web of problems for the right to be forgotten

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. K. Sijtsma, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op dinsdag 6 oktober 2020 om 16.00 uur

door

Paulan Eva Isolde Korenhof,

geboren te Alkemade

promotores:	prof. dr. R.E. Leenes prof. dr. E.J. Koops
leden promotiecommissie:	prof. mr. dr. M. Hildebrandt prof. dr. J.V.J. van Hoboken prof. mr. E.M.L. Moerel prof. dr. G. Sartor dr. M.L. Jones dr. A.P. Schouten

This project was funded by Stichting Internet Domeinregistratie Nederland (SIDN).

Printed by: GVO drukkers & vormgevers B.V.

Opdat we ons de waardevolle dingen herinneren.

Contents

Preface

Before you lies the end product of long hard work, sleepless nights, and litres of coffee. The journey up to here was a roller coaster with a bunch of ups and downs — sometimes in the same circle. I have learned a lot during this time, for which I am grateful to many people.

First and foremost, I would like to express my gratitude to my two supervisors, Ronald Leenes and Bert-Jaap Koops. Without them, I would have never been able to finish this journey. They guided me along the way, challenged my ideas, and shared their knowledge with me. I thank them for their tremendous support, invaluable guidance, and their patience.

A special, and sad, thank you goes to Sam Adams, my daily supervisor, who unfortunately died two years ago. I wish I could celebrate this with you.

I would like to thank all my colleagues at Privacy and Identity Lab (PILab), the department formerly known as the Tilburg Institute for Law, Technology, and Society (TILT), and Digital Security at the Radboud University for the inspiring working environment and interesting discussions. The discussions often kept lingering in my mind and provided me with the much needed inspiration to solve issues regarding either the process or the content of this dissertation. I would like to thank Mireille Hildebrandt for the philosophical input and talks we had during my time in Nijmegen, and Greg Alpar, for the thought provoking discussions. In particular I would like to thank Merel Koning, who started out as my office roommate at the Radboud and became a close friend. Without your support this dissertation would not have been what it is today.

Also, I would like to thank the members of the 'Timing the right to be forgotten'-panel at CPDP 2014, and my later co-authors: Jef Ausloos, Meg Jones, Giovanni Sartor, and Ivan Szekely. The inspiring discussion helped me in shaping a big part of my thoughts on the subject.

Thank you Nic, Michael, Sue, and my father, for helping me out by proofreading my chapters. Patrice, hf, Martino, LAG, revo, and the APV-people for inspiring discussions, the sharing of technical know-how, and out of the box input. Sam, Joan, Janneke, and Marlous, for being amazing friends who supported me through everything and often provided me with indispensable advice, distraction, or a trip to the sauna. The Café Libertad Kollektiv, for the writing juice. Ufomammut, Bongripper, and Godspeed You! Black Emperor for their music that helped me concentrate when I was too easily distracted or was surrounded by noise. And

of course, a big thank you to everyone I forgot to mention, but who helped me during this journey, supported me, provided me with feedback, inspiration, food for thought, coffee, or on occasion a direly needed Gin-tonic.

A special thank you goes to my parents for all the support, care, tremendous patience, and the many, many, peptalks. And last but not least, I would like to thank my partner(d), Ludo, for relentlessly bashing my analytical part, turning my life upside down, adding a lot of Oxford commas, and showing me how beautiful life can actually be. And the commodore 64.

List of abbreviations

AEPD	Spanish Data Protection Authority (Agencia Española de Protección de Datos)
AG	Advocate General
AJAX	Asynchronous JavaScript And XML
API	application program interface
CJEU	Court of Justice of the European Union
CMS/WCMS	Web Content Management System
CNIL	French Data Protection Authority (Commission nationale de l'informatique et des libertés)
DIK	Data-Information-Knowledge
DNS	Domain Name System
DPA	Data Protection Authority
DPD	Data Protection Directive
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
GDPR	General Data Protection Regulation
gTLDs	generic top-level domains
GUI	graphical user interface
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IT	information technology
LGBTQ	Lesbian Gay Bisexual Transgender Queer
SNS	social network sites
URL	Uniform Resource Locator
USB	Universal Serial Bus
WP 29	Article 29 Working Party
WYSIWYG	"what you see is what you get"-editor

Chapter 1

Introduction

Contents

1.1 The Web "never forgets"

We're an information economy. They teach you that in school. What they don't tell you is that it's impossible to move, to live, to operate at any level without leaving traces, bits, seemingly meaningless fragments of personal information. Fragments that can be retrieved, amplified ...

William Gibson, *Johnny Mnemonic*, 1981

The quote above seems like it came from a digital native, someone who grew up in the digital age, but it did not. Instead, it is an excerpt from *Johnny Mnemonic*, a dystopian science-fiction novel written by William Gibson in 1981. Gibson's science-fiction novels show a strong forecasting power with regard to future technologically driven developments. While we do not (yet) live in a world of holograms and human exoskeletons (both of which do seem to be on their way), we do already live in an information economy.

Information is of vital importance to us: the sharing and preservation of information is the key to our language, culture, science, society, and knowledge. We need information 'to live effectively' (Wiener, 1954, p. 17-18). Also on an individual level, the importance of information can hardly be overestimated: it forms the ground on which we base our choices and our understanding of the world and others around us. Given this importance of information, it is hardly surprising that there is an ongoing development of technologies that aid us in being better, faster, or more efficient with our information collection, processing, and analysis. Currently, Western society is heavily intertwined with, and dependent on, information technology (hereafter: IT). These technologies are the main vehicle for our communication, information access, and organisation of society. Even more, information fuels a part of our economic system and has even been called 'the new oil'[1]. The impact of IT on contemporary human life in total is so profound that Floridi even called it 'the fourth revolution' (cf. Floridi, 2014).

However, the blooming of the information age, also brought worries. Not only can information about governments, society, history, and culture be retained and made accessible through IT, but also with regard to personal information of private citizens there is an increase in the use of information technologies to record, store, adjust, and transmit this information. This increasing power of IT has led to the fear that development in IT will result in a world that does not forget and thereby traps us in an inescapable past (Dodge & Kitchin, 2007; Mayer-Schönberger, 2009; de Andrade, 2014; Burkell, 2016). One of the currently leading technologies, the internet, and more specifically, the World Wide Web (hereafter: the Web), especially fuelled these fears. The Web brought us world wide access to information

[1] See e.g., "The world's most valuable resource is no longer oil, but data", *The Economist*, 2017. `https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data`, last accessed 06-01-2019.

at the tips of our fingers. For most of contemporary Western life, it is even hard to imagine going about on a regular day without accessing the Web for something or the other. As we upload, display our preferences, search, write blogs, give our opinions, chatter on fora, join communities, maintain contacts, watch films, consult archives, complain about public transport, share a funny picture of our best friend, or show off our cats, we leave myriad snippets of personal information about ourselves and others on the Web. Some of this personal information could have unintended side effects and cause havoc. Yet, the fear is that once something is online, it will always be online. We can see this expressed by Rosen in his article with the telling title "the Web Means the End Of Forgetting" (Rosen, 2010). As such, the Web would lead to a world where we would consistently be reminded of and defined by that one online post that we cannot get rid of.

Sharing these concerns, several scholars argued in favour of the development of something along the lines of a 'right to be forgotten' (Mayer-Schönberger, 2009; de Andrade, 2014). The European Council concurred and aimed to address the concerns in the European Union General Data Protection Regulation (GDPR) by developing what has become known as the 'right to be forgotten', article 17 of the GDPR. Art. 17 GDPR should provide a counterbalance to the digital availability of personal information (Mitrou & Karyda, 2012). The article, dubbed in its last version as the "Right to erasure ('right to be forgotten')", gives individuals the right to obtain the erasure of personal information relating to them (art. 17(1) GDPR).

So, we had a problem, and now we have a means to solve it. Problem solved, right? Unfortunately, things did not turn out to be that easy with regard to art. 17 GDPR.

1.2 Houston, we have a solution!

Despite good intentions, art. 17 GDPR has had a rocky start. Ever since its announcement by Reding, art. 17 GDPR met with critique, scepticism, dread, and even outrage. Because the right allows individuals to have content relating to them erased, it raises concerns with regard to the freedom of expression and information (Fazlioglu, 2013; Larson III, 2013; Rustad & Kulevska, 2014; Kulk & Borgesius, 2018). On several occasions — especially in the media — the 'right to be forgotten' has been labelled as censorship and as a right that allows people to rewrite history.[2] It has even been called "the biggest threat to free speech on the

[2]See e.g., Adam Thierer, "Europe's 'Right to Be Forgotten': Privacy as Internet Censorship", *The Technology liberation front*, 2012. `https://techliberation.com/2012/01/23/europes-right-to-be-forgotten-privacy-as-internet-censorship/`, last accessed 24-10-2018; Jonathan Zittrain, "Don't Force Google to 'Forget'", *The New York Times*, 2014. `https://www.nytimes.com/2014/05/15/opinion/dont-force-google-to-forget.html?_r=0`, last accessed 30-03-2019; Jamie Grierson, "'Right to be forgotten' claimant wants to rewrite history, says Google", *The Guardian*, 2018. `https://www.theguardian.com/technology/2018/feb/27/right-to-be-forgotten-claimant-wants-to-rewrite-history-says-google`, last accessed 23-11-2018; Danny Sullivan, "Google Agrees To Complicated Worldwide 'Right To Be Forgotten' Censorship Plan" *Search Engine Land*, 2016. `https://searchengineland.com/google-to-`

Internet in the coming decade" (Rosen, 2011). Additionally, the framing of the right as a right 'to be forgotten' has been a topic of critique by opponents and proponents of the right alike, because this phrasing is taken to be misleading or even plain wrong (see e.g., van Hoboken, 2011; Koops, 2011; Powles & Floridi, 2014; Markou, 2015). I will discuss this particular point later in this study, once I have clarified other obstacles that need to be addressed first.

So far, the debate surrounding art. 17 GDPR is ongoing and includes inter alia questions on the practical application of the right in information systems and practices (see e.g., Politou *et al.*, 2018a; Sarkar *et al.*, 2018), the implications of the right for archives and historical research (see e.g., Szekely, 2014; De Baets, 2016; Vavra, 2018), the impact of the right on non-EU countries (see e.g., McDonald, 2019; Zeller *et al.*, 2019), how the right should relate to the passing of time (Ambrose, 2012; Sartor, 2015; Korenhof *et al.*, 2015), and how the right relates to other rights (see e.g., Li, 2018; Kulk & Borgesius, 2018). The debate includes many often contradictory views and opinions, and has the tendency to invoke "emotional and instinctive reactions (...) rather than rational and thought-through responses" (Bernal, 2011). While the debate is dense and all over the place, there is an extensive and ongoing discussion on what kind of right art. 17 GDPR is and what it should do (see e.g., de Terwangne, 2014; Bolton III, 2014; Bunn, 2015; Jones, 2018; Ausloos, 2018). For example, authors differ in their views on whether art. 17 GDPR is a new right (see e.g., Iglezakis, 2016), or an already existing right, albeit with some changes and in a new jacket (see e.g., Bunn, 2015). Another perspective to this is offered by Jones (previously: Ambrose) and Ausloos, who argue that art. 17 GDPR is a conflation of two different rights, namely of a 'right to be forgotten' that is related to the older French *droit à l'oublie*, and a more mechanic 'right to erasure' that ties in to the erasure of information as provided for by art. 12(b) of the (now outdated) Data Protection Directive (DPD) (Ambrose & Ausloos, 2013).

Moreover, the character of the right itself is also a topic of discussion. The right (either in its development phase or in its final version) has been labelled as or associated with a right to identity, to privacy, to be forgotten, to forget, to erasure, to deletion, to rehabilitation, to delisting, to obscurity, to cyber-oblivion, and as a right to be forgiven (see e.g., de Andrade, 2014; Ambrose & Ausloos, 2013; Xanthoulis, 2013; Hoffman *et al.*, 2015; Burkell, 2016; Voss & Castets-Renard, 2015). One of the most notable conceptualisations of a right to erasure, and specifically in the context of a 'right to be forgotten', stems from de Andrade, who conceptualises it as a "right to be different from oneself, namely one's past self" (de Andrade, 2014, p. 69). In this guise, ideally, the right should help individuals to develop themselves over time without having to fear from systematic stigmatisation of themselves in the here and now by their past actions and opinions (de Andrade, 2014). Along similar lines, but with a more explicit focus on the impact of technology and taking into account the 'life cycle' of information, we find Jones' analysis of the right to be forgotten as a way to realise digital oblivion (Ambrose, 2013). In this analysis, Jones ties the right to be forgotten to forgiveness

`censor-worldwide-sorta-243938`, last accessed 23-11-2018.

and states: "If the Internet Age will limit our ability to forget, it will in turn limit our ability to forgive or be forgiven" (Ambrose, 2013, p. 65). A right that allows under certain circumstances the erasure of information would be able to safeguard the realisation of forgiveness and "could help maximize the expressive potential of the Internet, while quelling anxiety related to an inhibited, exposed existence" (Ambrose, 2013, p. 75). However, not all conceptualisations of the 'right to be forgotten' tie the right's purpose and functionality to some form or function of forgetting. As already pointed out above, quite some authors explicitly argue against the conceptualisation of a right to have personal information erased as a 'right to be forgotten'. An example of a conceptualisation of the right that seeks to detach the right from the 'forgetting-framework' is offered by Bernal (2011). Bernal argues that the increasing collection of online available personal information is vulnerable to misuse, which in turn poses a threat to individuals and their autonomy Bernal (2011). Starting from the 'right to be forgotten', Bernal moves on to suggest recasting and renaming the right in order to address these issues. He suggests to introduce a right that builds on the idea that deletion should be the default in information processing and that an ongoing retention of personal information requires justification. With this functionality, combined with the emotional responses and misconceptions that the name 'right to be forgotten' evokes, Bernal argues that the right should be renamed 'right to delete' (Bernal, 2011).

The right to erasure and/or to be forgotten remains a topic of research and discussion up to present day and with scholars (re)analysing the right as well as the arguments used thus far in the debate (see e.g., Tavani, 2018; Jones, 2018; Ausloos, 2018). Not only are there many questions surrounding the how and what of art. 17 GDPR as a solution, but also the problems that it should address remain underexposed. With the core of the debate focused on the right itself, the character of the problems received little attention. Even more, cases that art. 17 GDPR was expected to resolve, turned out on closer inspection to be unsolvable or difficult to fully address with the right (Korenhof & Koops, 2013; Korenhof, 2014). Exemplary for this is the case of the 'Drunken Pirate'. This case received much media attention and has been used to illustrate the problems of the memory of the Web (Mayer-Schönberger, 2009; Rosen, 2010). In this case a young woman, referred to here as 'S', experienced first hand how just a bit of personal information on the Web can affect your life in a destructive manner. An online photo showing S with a pirate hat drinking from a plastic cup, captioned 'drunken pirate', played a role in her failing her internship and thereby ending her career as a teacher. Mayer-Schönberger writes: *"S[...] considered taking the photo offline. But the damage was done. Her page had been catalogued by search engines, and her photo archived by Web crawlers. The Internet remembered what S[...] wanted to have forgotten."* (Mayer-Schönberger, 2009, p. 1).

A closer inspection of the case showed that the core of the problem (at least of the online part of the problem) in this case was caused by the fact that an unintended audience managed to get access to the content of S's MySpace; a colleague of S accessed the content and a few days later S was told that she failed

her internship, partially due to the content that was viewed on her MySpace. The Web thus played a certain role in the chain of events, but not by providing an eternal memory. Instead, the technology allowed S to accidentally reveal the photo to unintended audiences. Assessing the mechanisms of art. 17 GDPR, it quickly turned out that art. 17 GDPR is unable to resolve the problem in this case.[3]

What the 'drunken pirate' case shows is that in order to understand if, and for what kind of cases art. 17 GDPR is a viable means to resolve the issue, we need to give more attention to the particular role that technology plays in the problems. However, the role of the technology can also be a matter of dispute. This is one of the issues shown by the widely discussed *Google Spain* case[4] (for the tip of the iceberg, see e.g., Kuner, 2014; Frantziou, 2014; Kulk & Borgesius, 2014; Cofone, 2015; Lynskey, 2015; De Hert & Papakonstantinou, 2015). It is a landmark case ruled by the Court of Justice of the European Union (hereafter: CJEU) and is often understood as a case where a 'right to be forgotten' is granted to an individual (see e.g., Wolf, 2014; Jones *et al.*, 2015; O'Hara, 2015; Post, 2017). The case is characterised by a dispute on the role of search engine technology with regard to the display of relatively old personal information.

The case revolves around two online search results that are displayed in response to a name query. When typing in the name of the plaintiff in the case, 'G', a Spanish citizen, Google Search displayed two links to newspaper articles in the Catalan newspaper *La Vanguardia* from 1998. The articles were originally published in the (printed) newspaper of 19 January 1998 and 9 March 1998 and were uploaded as a part of digitising *La Vanguardia's* archive.[5] The articles concisely announced the real-estate auction connected to social security debts of G. The newspaper was legally obliged to print the information on behalf of the Spanish Ministry of Labour and Social Affairs.[6] G requested *La Vanguadia* to remove the articles as well as Google to remove the links connecting these articles to his name. Both *La Vanguardia* and Google refused. In 2010 G filed a complaint at the Spanish Data Protection Authority (the *Agencia Española de Protección de Datos*, AEPD) in the hope of resolving the issue. The AEPD dismissed the claim with regard to *La Vanguardia*, because *La Vanguardia* was legally obliged to publish the information. However, in reference to Google the AEPD upheld G's claim and ordered Google to remove the links. Google disagreed with the decision of the AEPD and brought the case before the National High Court of Spain to fight the AEPD decision. In turn the National High Court of Spain requested a preliminary ruling from the CJEU on the case. One of the main legal

[3]For a full discussion of the 'Drunken Pirate' case and an in-depth explanation of why this is not a case for art. 17 GDPR, I would like to refer the reader to my paper *Stage Ahoy! Deconstruction of the 'Drunken Pirate' Case in the Light of Impression Management* (Korenhof, 2014).

[4]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[5]http://www.lavanguardia.com/hemeroteca, last accessed 20-08-2017.

[6]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §16.

questions in this case is whether a search engine can be considered a controller of the information that it displays (see e.g., Stute, 2014; Lindsay, 2014; Lynskey, 2015; Cofone, 2015). The answer to this question depends on whether one holds the view that a search engine operator determines the purposes and means in which a search engines processes personal information in search results — thus falls under the definition of *controller* in accordance with art. 2(d) DPD. At the core of this question lies the more fundamental question: what is a search engine, and what does it do?

The *Google Spain* case shows a fundamental difference in views on the role of a search engine: the interpretation of search engines by Avocate General Jääskinen (hereafter: AG) who advised the CJEU in the case, and the views of the CJEU stand in stark contrast to each other. On the one hand, there is the AG who takes search engines to be a neutral and truthful intermediary that sets up "automated, technical and passive relationships to electronically stored or transmitted content"[7] over which the search engine operator has no control. On the other hand, there is the CJEU who argues that a search engine operator is a controller, because a search engine performs actions "additional to that carried out by publishers of Websites"[8]. As such, the search engine operator determines the purposes and means of these activities — and thus in turn processes the information that is indexed from other websites.[9] This difference in views is also visible in the literature surrounding the *Google Spain* case with, on the one hand, authors who argue that search engine operators should not be considered the data controller of the search results (see e.g., Sartor, 2014; Peguera, 2015), and on the other hand, those who argue that they should be considered as such (see e.g., Hijmans, 2014; De Hert & Papakonstantinou, 2015). The core difference between these views boils down to the question whether the search engine 'does' something to the information when displaying it as a search result. In the end, the CJEU ruled that G. had the right to have the search results removed, while the original content on *La Vanguadia* remained untouched. The ruling (and following wave of erasure requests) was met with general unhappiness by a significant part of the legal and IT professional community, and gave rise to — again — an ongoing discussion, this time about how and when to apply erasure to search results, and how to balance these with other rights (see e.g., Singleton, 2015; Kampmark, 2015; Bougiakiotis, 2016; Youm & Park, 2016; de Mars & O'Callaghan, 2016).

The road to the introduction of art. 17 GDPR has thus been paved with unclarity and disagreement that has plumbed deep into the core of the right — and the discussion continues (see e.g., Ranquet, 2019; Padova, 2019; Yaish, 2019). A critical cause of the right's problems can be traced back to uncertainty and vagueness with regard to the exact manner in which online technology can cause problems for individuals by being used to process their personal information.

[7] Opinion Advocate General Jääskinen, 25-06-2013, C-131/12, ECLI:EU:C:2013:424 (*Google Spain SL, Google Inc./AEPD, G*), §87.

[8] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §35.

[9] Ibid., §33.

1.2.1 Forty-two: and now what?

"All right," said Deep Thought. "The Answer to the Great Question..."
"Yes..!"
"Of Life, the Universe and Everything..." said Deep Thought.
"Yes...!"
"Is..." said Deep Thought, and paused.
"Yes...!"
"Is..."
"Yes...!!!...?"
"Forty-two," said Deep Thought, with infinite majesty and calm.

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, 1996

Art. 17 GDPR somewhat resembles 'forty-two', the answer given by supercomputer Deep Thought to the ultimate question of life, the universe, and everything, in Douglas Adams' famous science fiction novel *The Hitchhiker's Guide to the Galaxy* (Adams, 1996). The problem with the answer, as Deep Thought states it, "is that you've never actually known what the question is" (Adams, 1996, p. 121). This seems to be the problem of art. 17 GDPR as well. While the right appeals to everyone's imagination, it is unclear for what kind of problems art. 17 GDPR actually is a suitable answer.

The goal of this research is to fill this gap by providing for a better understanding of the kind of issues that art. 17 GDPR can resolve, what kind of issues it cannot resolve, and how art. 17 GDPR can best be applied. In order to analyse what and how, and even if, art. 17 GDPR is viable as a means to address the issues at hand, we need to understand *what* the problem is, and *how* it comes into being. In order to say anything about the viability of art. 17 GDPR to address problems, it is necessary to first understand the problems. Yet, how do you find the problem to which art. 17 GDPR is the answer?

In *The Hitchhiker's Guide to the Galaxy*, Deep Thought designs another computer to calculate the question, and this computer is Earth. Luckily, we already have Earth. And it is in this place where problems can be found, if any, to which art. 17 GDPR is the answer. The identification and understanding of the problems and the manner in which they are brought about, therefore forms a major part of this study. However, because 'Earth' could in theory cover a scope of life, the universe, and everything, I need to reduce the scope of the problem identification to workable dimensions. This places some issues outside the scope of this research, but these can be topics for future research. I set up the research area with a set of parameters, namely 1) personal information, 2) the Web, 3) availability of the information for common users, and 4) the problematic character of the information processing. I explain these parameters and the reason for picking these below.

1. Personal information The parameter 'personal information' is an obvious one. With this parameter, I follow the material scope of the GDPR which sees to the protection of "any information relating to an identified or identifiable natural person" (article 4 (1) GDPR). However, for the purpose of this study, I restrict my focus to a particular subset of personal information. I have chosen to rule out issues that generally already are covered by legislation, like identity theft and libel, and instead focus on information that is — at least initially — not the subject of an offence.

2. The Web The GDPR concerns the processing of personal information. This can cover many sorts of information technologies. For this study, I have chosen to focus on information processing on the Web. The reason for this is threefold. First of all, due to the popularity and public role of the Web, I take the problems in this environment, as well as the interest in a good application of art. 17 GDPR, to be particularly relevant. Secondly, most of the discussion surrounding art. 17 GDPR is focused on its online application. I therefore assume that it is valuable to help clarify the discussion with regard to these forms of information processing. Lastly, the current use of the Web is one of the driving forces for the development for art. 17 GDPR.[10]

3. Availability for users In order to restrict the scope of my research to a workable dimension, I have chosen to focus on a particular part — or rather presentation — of online information processing, namely the public and semipublic presentation of information to common Web users. With users I refer to common civilian — not necessarily civil — users of the Web in the broad sense of the word. These users can be natural persons, but also companies, employers, professionals, etc. The core describing criterion for the user group that I focus on, is that they interact as a human being directly with the front end of the Web.

I have chosen this user focused angle, because so far user access has been the main focal point in the art. 17 GDPR debate and cases; e.g., both the drunken pirate case and the *Google Spain* case revolve around issues that relate to the accessibility of online information for regular users. The focus on public and semipublic online content is to tie in to the main functionality of the Web as an open communication network. The combination of these two focal points leads me to concentrate on the manner in which online personal information is (semi)publicly presented to users as a result of the Web's technological architecture, design choices and user actions. With this angle, I place the emphasis of my research on the manner in which the Web presents information to the perception of the user. However, in order to make sense of how this presentation came to be, I cannot stick merely to what is available to the perception of users, but also need to take a peek 'under the hood' of the Web and look at its information flows. The relation that is at the centre of this research is thus: user $\leftrightarrow$ Web $\leftrightarrow$ individual.

[10] See Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*, 2012. `http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

It is important to note that the restriction of the research to what is available and presented to common Web users places this study's focus on the the user-side of the Web — the front end. This restriction cuts off the investigation of another set of problems of online personal information, namely the use of personal information by corporations and the like not for public or semipublic online presentation purposes of that information, but for profiling, trading, risk analysis, service improvement, etc. This takes place behind the screens and generally entails the aggregation and further processing of personal content on a massive scale. While I certainly regard this as an important problem area, the research thereof would require a different research angle if I want to do justice to the impact and complexity of the problems. Given the fact that, so far, the debate surrounding art. 17 GDPR has had its focus on the front end of Web use, I gave priority to researching the manner in which Web content is presented to users over the implications of back end processing practices.

4. Problematic character I am looking for problems that are raised by the presentation of online personal information to Web users. The focus here lies on the problems raised by the information processing itself, and not on the problematic consequences that result from the presentation of this information to users. With the consequences of the presentation of the information, I mean the actions that users undertake based on the information they encounter. To give an example, take a case where Harry is dating Sally and stumbles online on a nasty comment that Sally has made about dogs. Harry finds the comment awful and decides to end the relationship. In this research, I concentrate on how this particular presentation of Sally reaches Harry or Sally herself and whether this is problematic in any way *in relation to* Sally. For example, if the content was posted by another Sally and seemed to be coming from Harry's Sally, or if the comment at the time was a joke, but lost its context over time, the processing establishes a problematic representation of Harry's Sally to Harry, and possibly also to Sally herself (e.g., the comment reminds her of a bad period in her life). The main point of attention is thus not how Harry decides to act based on his encounter with this online comment or how Sally reflects upon herself (although I touch upon such consequences of the representation of an individual occasionally), but the manner in which Sally is presented to Web users like Harry and herself. The reason for this focus, is that the GDPR sees to information processing, and not on addressing human responses. I therefore take the core of the problems that art. 17 GDPR aims to resolve to lie in the processing of personal information and how it symbolises and presents people to Web users. The researched problems are thus of a symbolising nature.

The main question and goal The goal of this study is to assess the merits of art. 17 GDPR to address problems raised by the presentation of online personal information to users. Given the fact that technological developments have been the leading reason for the development of the GDPR, and particularly art. 17 GDPR, a significant part of this study is focused on exposing the manner in which the technological constitution of online information sources affects the manner

in which personal information is available and presented to users. Only when this is clear, can we turn our attention to the assessment of the functionality and viability of art. 17 GDPR to address these problems. The main research question is therefore comprised of two core parts: (1) the problems raised by the presentation of online processed personal information to Web users, and (2) art. 17 GDPR as a viable means to address these problems. It is formulated as follows:

To what extent is art. 17 GDPR a viable means to address problems for individuals raised by the presentation of online personal information to Web users?

By answering this question, I aim to give people who deal with online presented personal information, and especially those who work on solving the problems that the availability of this information may cause, more grip on the problems, as well as an idea of what art. 17 GDPR can do in these cases. I therefore take the main audience to be controllers of online media, IT specialists, lawyers, and policy makers, who have an interest in getting an in-depth view of the character of the problems raised by the presentation of online personal information to Web users and the viability of art. 17 GDPR to address them.

1.3 Methodology

The main question of this study requires an investigation of the extent to which a legal tool is able to resolve issues that result from the manner in which online processed information can affect our relation to and interaction with personal information. To properly answer the main research question, this investigation consists of two parts: an analysis of the problems, and an analysis of the proposed legal means to address these problems. As such, the research topic lies at the crossroad of various elements; namely human beings, technology, information and law. In order to identify the problems, the research needs to be both exploratory and explanatory. The main question requires me to explore and explain how the Web affects the relation between users and online personal information. It therefore cannot be answered by doing only legal research. I also need to delve into the impact of technology on the relation between personal information and human beings. Answering the main question therefore necessarily requires an interdisciplinary study, because I need to get some grip on how the Web works, how humans behave, as well as how law should be applied. However, different disciplines have a different understanding of concepts (an example of this is the various takes on 'information', which I will discuss in the next chapter). This especially seems to be the case with regard to lawyers and technicians. This study therefore was not a case of just conducting a research, but also a journey back and forth between different disciplines to identify the problems and to find the right language to confer the message.

Not only trained as a lawyer but also as a philosopher, I have used this part of

my background as glue to connect the various views and elements. The best methodological 'glue' that allows me to take all these various disciplines and elements into account and assess the impact of the Web on the presentation of personal information to users, is to use a *postphenomenological* approach for this research. Postphenomenology (like its 'mother' phenomenology) is focused on the manner in which human beings experience the world. It does this with "a starting point in empirical analyses of actual technologies"(Rosenberger & Verbeek, 2015, p. 9). As such, postphenomenology "combines an empirical openness for the details of human-technology relations with phenomenological conceptualisation (Rosenberger & Verbeek, 2015, p. 32).

Postphenomenological research has several main characteristics:

1. Focus on human-world relations Postphenomenology studies technology "in terms of the *relations between human beings and technological artifacts*, focusing on the various ways in which technologies help to shape relations between human beings and the world" [emphasis original](Rosenberger & Verbeek, 2015, p. 9). Technology is approached as something that *mediates* the human experience with the world (Rosenberger & Verbeek, 2015, p. 11).

2. Empirical investigation combined with philosophical analysis In its study of the relation between human beings, technology and the world, postphenomenology takes in a pragmatic angle where it is closely tied to technology as well as the human praxis. Rosenberger and Verbeek state: "*In order to understand a technology or a technological development, postphenomenology always analyzes the character of the relation human beings have with this technology and the ways in which it organizes relations between human beings and the world*" [emphasis original] (Rosenberger & Verbeek, 2015, p. 13). For this, postphenomenology combines "*philosophical analysis with empirical investigation*" [emphasis original] (Rosenberger & Verbeek, 2015, p. 9). Postphenomenology can therefore be described as 'empirical philosophy' (Rosenberger & Verbeek, 2015, p. 30). The empirical evaluation of technologies and their corresponding practices are the starting point for the philosophical investigation and reflection: "In order to understand human-technology relations, an empirical account is required of the role actual technologies play in human experiences and practices" (Rosenberger & Verbeek, 2015, p. 31). This empirical part of the research can be based on "empirical work by others, from self-conducted studies, or from an analysis of first person experiences that involve specific technologies" (Rosenberger & Verbeek, 2015, p. 17). Due to its strong empirical outlook, "[a]n essential aspect of the postphenomenological perspective is its focus on case studies of concrete human-technology relations" (Rosenberger & Verbeek, 2015, p. 32). As I explain later, I will account for this empirical angle by making use of the work of others, as well of my own first person experiences.

3. Constitution of the subject and the object "*[P]ostphenomenological studies typically investigate how, in the relations that arise around a technology, a specific 'world' is constituted, as well as a specific 'subject'*" [emphasis original] (Rosenberger & Verbeek, 2015, p. 31). In this relation, the 'world' is constituted as a particular object by the technology for a certain perceiving human subject (Rosenberger & Verbeek, 2015, p. 31). The world in this context is that which is outside of the perceiving human being and is the object towards which her attention is directed. For example, by using a microscope, someone focuses on 'the world' through this particular technology that makes bacteria and the like visible to the person using the microscope. With this, a particular framing of the world (bacteria on a microscopic level) is presented as an object to the subject using the microscope.

4. Conceptual analysis On the basis of the three elements above, "*postphenomenological studies typically make a conceptual analysis of the implications of technologies* for one or more specific dimensions of human-world relations" [emphasis original] (Rosenberger & Verbeek, 2015, p. 31). The goal of the postphenomenological study of the manner in which technology affects the human-world relation is to identify the implications that the technology in this role has for the human subjectivity as well as for the object that is the focus of the human agent (Rosenberger & Verbeek, 2015, p. 9). The identification of the implications generally is focused on one or more specific dimensions of the relation between the user and the outside world (Rosenberger & Verbeek, 2015, p. 31).

Due to its detailed focus on actual technologies and their impact on the relation between human beings and their world, postphenomenology is a suitable qualitative research methodology to gain an in-depth insight into the mechanisms at play and the implications that may result from the manner in which the Web affects the relation between human beings and personal information. However, applying a postphenomenological method does require quite some methodological choices from the researcher, because next to these mentioned main characteristics, there is not one clearly defined 'postphenomenological method'. There is much diversity in the methodology that postphenomenologists use (Rosenberger & Verbeek, 2015, p. 10). I will therefore discuss here how I chose to implement the main characteristics of postphenomenology in order to conduct this study.

1.3.1 Focus on human-information-world-relation

In this study I focus on the relation that the Web brings about between users and their view on a particular individual by providing these users with information. For the purposes of this study it is therefore not only important to look at the relation that the Web establishes between users and the world, but specifically to look at the manner in which information takes shape in this process. The focus is therefore on the relation between real life individuals and their online representation for the perception of Web users.

In this relation, the Web takes on a *mediating* position. The view on technology as mediating relations between humans and their world is the core of postphenomenology and will be discussed in detail in chapter 3. However, due to the role of information in this, it is necessary to combine a theory of information and interpretation with postphenomenology. Before going into the mediation theory, I therefore first delve into the relation between the world, information and human beings (world → information → human being) in chapter 2. By combining these two, I employ a hermeneutic postphenomenological approach to address the first part of the main research question, the problem analysis.

1.3.2 Combination of an empirical investigation with philosophical analysis

In order to determine whether and how art. 17 GDPR can resolve problems with personal information on the Web, we need to understand how the Web's technology mediates this information to users. The starting point of the first part of this study is therefore the technology of the Web. This technology is empirically investigated and combined with philosophical analysis. This research is based on first person experience, participatory observation, and on work done by others. The work of others on which I build my findings, originates from social sciences, computer sciences, media theory, philosophy, as well as on reports and guidelines provided by technologically oriented groups and organisations like the W3C and the Internet Engineering Task Force. For the induction of some general effects and experiences that result from the way in which the Web influences our relation to personal information, I depend on my own perception to cross-refer certain elements. This personal character of this perception does place a particular stamp on the research; had I for instance been visually impaired and relied more on the audio related applications on the Web, the reflection of the manner in which the Web mediates personal information would take a different shape at some points.

A significant portion of this study entails a micro-level analysis of technology, which is typical for postphenomenology. By a micro-level analysis of the technology, I mean the analysis of detailed concrete user mechanisms on the level of the perception of the individual user (for example, how a user can use a search bar in a browser). However, restricting myself to micro-level analysis would be problematic, because it would exclude certain macro dimensions, like the business models underlying the technology, that are also of importance to get an understanding of the way in which the Web affects the user's relation to personal information. Moreover, in order to identify and clarify how the problems emerge, I cannot restrict myself to what is directly experienced by users. Instead, I also need to look at what information technology does not show us. To investigate *how* the problems arise in the processes of the technological event, I also need to look at hidden social and technological mechanisms that gave rise to particular presentations of information. I therefore also at points incorporate a macro-level analysis. While the core approach of postphenomenology lies in a micro-level analysis, it is flexible enough in its methodology to allow combination with macro-

level analysis, as well as other approaches.[11]

My point of departure is — in the broad sense of the word — a case study. A case study is "an empirical inquiry that [...] investigates a contemporary phenomenon within its real-life context, especially when [...] the boundaries between phenomenon and context are not clearly evident" (Yin, 2009, p. 49). A case study is a valuable method if the goal of the research is to explore the real-life context and conditions of a particular phenomenon (Yin, 2009, p. 49). Case studies are therefore good for answering 'how' and 'why' questions when the researcher has little or no control over the environment (Yin, 2009, p. 29). These questions and the conditions are applicable to my main question. In order to identify what exactly the problems are, we need to know *why* they are problems and *how* they are caused. Additionally, I have little or no control over the Web as an environment. Even more importantly though, in order to get a strong grip on the character of the problems, an analysis of the manner in which the Web mediates the interaction between users and personal information is needed in its full-coloured context of daily life. A case study allows me to investigate such complex phenomena, where there are many variables at play, as well as multiple sources of evidence. This helps me to identify the general character of the problems and the elements that play a role in how they come to be. However, a remark is in order here. The case at the heart of this study is the presentation of personal information to users by the Web. This is not a single case, but a substantial range of (possible) cases. As I want to identify not only existing, but also potentially future problems, I decided to not restrict my research to a handful of concrete cases of user-technology interaction that raised problems, but instead trace the general mechanisms of this interaction. I therefore take a broad approach to the case study.

Additionally, the way in which personal information is made accessible to users on the Web is too extensive to research in one go. The Web is not a single technology and covers miscellaneous aspects and sub-technologies. This is why I decided to split the research up into several 'sub-case' studies. These sub-cases are several online applications that often deal with public or semipublic personal information, and one online phenomenon. The choices for these sub-cases are based on what I perceived to be a logical split up of Web applications complemented with the cases that surfaced in art. 17 GDPR debate. The cases see to 'regular' web pages, social network websites and search engines. These three seem to cover most of the Web applications, although possibly in a hybrid form. However, the perceptive reader may have noticed that next to writing about online applications, I also referred to 'one phenomenon'. When exploring individual cases like the Technoviking and the Star Wars Kid case, I found that these cases were

[11] I gather this from Keymolen's presentation in the *Postphenomenology & Politics* panel on 12-07-2019 at the *Human-Technology Relations* conference in Enschede, the Netherlands, as well as from Rosenberger's closing session on 13-07-2019 at this same conference. Additionally, in this context it is relevant to refer to the article *Technological Mediation and Power: Postphenomenology, Critical Theory, and Autonomist Marxism* by Rao *et al.* (2015). In this article, the authors argue that we can best understand postphenomenology and critical theory as complementary of each other.

not confined to a particular, or an interplay of two particular online application(s). Despite having seemingly similar problem mechanisms, some cases stretched over various kinds of online applications; these are the 'viral' cases. Because online virality does not fit any particular sub-case, but it does fit within the main case as a specific phenomenon, I dedicate a separate chapter to it.

1.3.3 Analysis of the implications for the constitution of the object

The inducement of this study is the potential problems that individuals may experience as a result of the online availability of information about them for Web users. The underlying mechanism at play here is that people use the information that they have (or think they know) to form a view of others as well as of themselves, and act based on this information. The Web affects the knowledge that users have about people by potentially offering personal information to them and presenting it in a certain context and manner. The main question is thus focused on the question of how the online personal information can affect a user's interpretation of the people (including herself) whose personal information she may encounter online. As such, the dimension that is being addressed here is symbolic (in the literal sense): how does the technology of the Web affect the way in which individuals are symbolised by their personal information?

The analysis of the implications that I derive from the cases, therefore has a very specific focus: at the centre of attention is the issue of how the Web mediates personal information towards users constitutes a certain view of an individual for these users. While at many points the constitution of the subject acting with the technology certainly plays a role in this, the main emphasis of this research is on understanding how the object is constituted *for* the subject. The focus thus lies mainly on the constitution of a specific object in the relation, namely the presentation of a particular individual by means of personal information.

The analysis of the implications form the first part of answering the main research question.

1.3.4 Conceptual analysis and evaluation

At the end of each (sub)case analysis chapter, I draw some general conclusions about the implications that result from a particular technology with regard to the manner in which personal information is presented to users, and how these can raise problems for individuals. However, at that point I am not done yet. In order to assess the capability of art. 17 GDPR to address these issues, I also need to examine how art. 17 GDPR works. I examine the workings of art. 17 GDPR and the rationale behind the provision. I leave aside its role in the bigger legal framework and corresponding legal complexities that may arise as a result of for instance its territorial scope. The reason for this zoomed-in approach is that I want to test the base functionality of the right to address the issues. If the right in its functional application is not capable of solving the problem (as was the case

with the 'Drunken pirate'), it will certainly not succeed when it gets entangled in broader issues (e.g., cross jurisdictional). The underlying idea is to provide the reader with an idea of the value of art. 17 GDPR as a viable means in itself for problems brought about by the manner in which the Web mediates the interaction between users and personal information.

To actually assess the viability of art. 17 GDPR to address problems, I combine the problem analyses of the sub-cases with my conclusions on the mechanisms of art. 17 GDPR. I construct an overarching view, in which I go more deeply into the conceptual analysis and propose a particular conceptualisation of art. 17 GDPR. This bigger picture in turn serves as the backdrop for the assessment of the viability of art. 17 GDPR to address the identified issues.

1.3.5 Scope

The goal of this study is to uncover the underlying issues of the manner in which personal information is presented to users online: what are the main mechanisms, and what is the cause of the problems? However, because the Web is a complex technology where multiple factors are at play, the question is: how far do I go? Every aspect that I examined seemed to lead to another. I went from software to hardware, to program languages, to bits, to transmission methods, to economic models, to cognitive theory, to neurons and to circuit boards, and so on. There was no end to it. The result is that the research turned into something much like an oil spill in the sea; while I started with an oil drop in one location — the Web's interface approached from a postphenomenological perspective — the oil soon spread over an increasingly broad surface, and I started to view the topic from more phenomenological, anthropological, pragmatic, economic, informational and empirical perspectives. The research slowly started to cover an ever-expanding surface of perspectives and topics that tie to the current one to such a degree that I risked conducting a study that would never be finished until every ocean, river, ditch, and puddle was covered. I therefore had to curb this research and had to leave some interesting trails unexplored. Some parts have been cut out to maintain a general focus and line of argument. An example of such a cutout, is the analysis of the impact on the Web and user interaction of the diverse protocols that give shape to the internet. The guiding criterion here, was that I wanted to maintain the focus on the composition of online applications from the user perspective. I therefore focused on the front end of these pages as displayed in a standard desktop Web browser. This meant, for instance, that I gave little attention to the aggregation of personal information in the back end of online services, as this would require a different kind of research. Additionally, I also gave little attention to the differences in the interaction with online content between various devices like desktops, smartphones, and tablets. While these devices play a fundamental role in our interactions with online content because they are necessary to realise our interaction with the Web, analysing their differences in mediation would shift the focus from the Web to the device and thereby away from the subject matter that lies at the heart of this study. The reader should therefore note that behind

and in combination with every examined mechanism and element, there are often multiple other mechanisms at play, and occasionally the reader's attention will be drawn to these unexplored trails in the footnotes.

1.3.6 Disclaimer with respect to theory limitations

The difficulty with law as well as philosophy is that there are so many possible nuances and exceptions, that trying to say something about a topic from one of these perspectives can at times paralyse the author in a realm of endless coverages against accusations of missed or overly simplified points. This turned out to be particularly crippling with regard to the scope of this study as there was no word limit and the topic touches upon law as well as philosophy. This has led to some versions of sections that were so extensive that the reader would easily lose sight of the main line of argumentation, that is, if she managed to overcome the boredom to get through the section altogether. I therefore have, at times, streamlined or simplified descriptions and analyses, and left the more remote exceptions or situations undiscussed.

Moreover, during the course of this study I came to the understanding that there is no such thing as a 'one size fits all' theory or discipline that could function as a complete framework to investigate the issues at hand. I therefore have to rely on several theories and disciplines to identify and explain distinct pieces of the puzzle. While most theories partially overlap, they do not exactly match. In many theories words are taken to have a specific meaning, which can have a dissimilar meaning in another (e.g., the scope and particularities of concepts like 'object', 'sign', and 'information' can vary highly per author). Adding to this is the complication of the use of neologisms by many authors that I discuss. In order to cope with this predicament, I had to smooth over differences and use a self-chosen and described terminology to connect the dots and prevent mismatches of the word use of the thinkers discussed. In some cases this does not do justice to the nuances and detailed character of the theory of the original thinker from which I draw inspiration (one of the examples being my use of the ideas of Charles Sanders Peirce). I regret this, but given the fact that I aim to keep the size of this study contained to one volume of moderate proportion, I see no way around it. Thus to prevent pages of nuances that add little to nothing to the main argument and extend the size of this dissertation with dozens of pages, let this serve as a disclaimer for the full extent of this study (albeit the lawyer as well as the philosopher in me are sometimes unable to help themselves and make some minor disclaimers throughout the text). I hope therefore, dear reader, that you can accept this and excuse me at points for being bold with certain theories, aligning some unconventional combinations, and simplifying or straightening out certain details and exceptions. I sincerely apologise to any experts whose toes may cringe during the reading of this dissertation, as well as any of the philosophers who turn in their graves at my 'remixing' (to use a term from the popular internet culture) of their original ideas.

1.4 Overview book

In this book, I research to what extent art. 17 GDPR can be seen as a viable means to address problems for individuals raised by the presentation of online personal information to users. A significant part of the research consists of an analysis of three applications of technology, and one online phenomenon in order to identify the character of the problems. I am interested in how in these cases the technology establishes a particular relation between human beings and information. However, because information, as well as technology, are complex things, I first need to get some general sense of what I am looking at. I therefore first pave the way for these analyses by discussing my theoretical framework in chapters 2 and 3. Following these chapters, I delve into the cases. After finishing the case analyses, it is time to turn my attention to art. 17 GDPR. To assess to what extent art. 17 GDPR can be used as a viable means, I first need to get a grip on how art. 17 GDPR works. This I examine in chapter 8. Next, I bring the previous chapters together in an overarching view of the problems, and assess which problems art. 17 GDPR can resolve. Finally, I conclude this study by answering the main question and assess the suitability of art. 17 GDPR as a means to address the problems that I identified.

The structure of the book is as follows:

Chapter 2: Framework part I: information and the informational persona In this chapter, Framework part 1, I discuss the concept of information and how human beings relate to information. I introduce the first part of the theoretical framework and the set of conceptual tools that I use in the rest of the book to examine how the Web raises problems by presenting personal information to users. The core concepts that I introduce in this chapter are the 'signifying object', the 'presence' of information, and the 'informational persona'.

Chapter 3: Framework part II: Technological mediation and personal information In the second Framework chapter, I delve into the non-neutral role of technology as signifying object, and the manner in which technology can make information present. The most important elements that I explain in this chapter, are the 'mediating' role of technology, and the 'intentionality' of the technology herein.

Chapter 4: Web pages I start the sub-case analysis by examining regular web pages in chapter 4. In this chapter, I focus on the manner in which information is encoded on web pages and made accessible to users. A central point of attention here is the impact that the digitisation and online encoding of information has for the relation between users and personal information. Because all the other sub-cases also see to content on the Web, and thus necessarily involve a certain kind of web page, chapter 4 also serves as the foundation for the following chapters.

Chapter 5: Social media In chapter 5, I examine the implications of social media use for personal information on the Web. I look at the impact that the main mechanisms that are typical for social media have for the processing of personal information online. For this, I base myself on the highly popular social media platform, Facebook.

Chapter 6: Search engines In chapter 6, I examine the impact of search engines on the relation between users and personal information. I focus my attention on the most used search engine, Google Search. I examine its implications for the presentation of personal information to users by looking inter alia into the search engine's ranking mechanisms as well as into its position on the Web as an information realm.

Chapter 7: Going viral In chapter 7, I explore the phenomenon of online virality, and the implications that it has for the presentation of personal information to users. For this, I look into several viral cases, as well as into research on virality in order to get an idea of the main mechanisms of virality and its impact on the presentation of personal information.

Chapter 8: Art. 17 GDPR In chapter 8, I focus fully on art. 17 GDPR. By means of close reading, and with the support of case law and the work of other legal researchers, I examine the mechanisms of art. 17 GDPR. Additionally, I pay attention to its names 'right to be forgotten' and 'right to erasure'. The findings of this chapter are used in the next chapter in combination with the case chapters to research what kind of problems art. 17 GDPR can resolve.

Chapter 9: Art. 17 GDPR and the problem narrative Chapter 9 is the heart of this study. Lending inspiration from the work of Ricoeur, I bring the previous chapters together in a bigger picture that provides us with an overarching problem analysis. This analysis is then used as a backdrop to assess to what extent art. 17 GDPR can address the identified problems.

Chapter 10: Conclusion In chapter 10, I conclude on the extent to which art. 17 GDPR is a viable means to address problems for individuals raised by the online processing of personal information and its availability to users.

Chapter 2

Framework part I: information and the informational persona

Contents

2.1 Introduction

The main research question of this study concerns the crossroads of the Web, an information technology, and the GDPR, which sees to the protection of personal information. Before I can start exploring the problems, I first need to get a grip on two things: 1) the relation between information and human beings, and 2) the role that technology plays in this relation. In this first framework chapter, I shall focus on point (1). The goal of this chapter is to introduce the theoretical toolkit and background needed to get a grip on what we are actually looking for when we turn our gaze towards technology and the manner in which it can affect our relation to personal information. In this chapter, I shall therefore delve into the manner in which we relate to information as perceiving party, how information relates to us as a subject of the information, and lastly what this means for interactions between humans agents. However, first, it is important to understand what 'information' is. This already poses me with a challenge because 'information' is a complex concept with no unified definition; the concept is defined from myriad perspectives and contexts by various scientific disciplines (cf. Losee, 1997; Floridi, 2011, 2016; Aamodt & Nygård, 1995). In, inter alia, information science, cybernetics and philosophy of information, numerous authors have attempted to provide a satisfactory definition of information. One of the more famous concepts of information is Wiener's view in which he poses the concept of 'information' as the opposite of 'entropy': "Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization; and the one is simply the negative of the other" (Wiener, 1961, p. 11).[1] Other authors place the core of the definition at other elements, like the process that generates information and/or the value attributed to the content. For example, Losee gives a more process oriented definition of 'information' by stating that information is "the values within the outcome of any process" (Losee, 1997, p. 254).

The various ways in which information is understood, are generally highly intertwined with the scientific discipline of the author defining the concept. Authors like Losee and Shannon therefore argue that most definitions of information define a subset of information that is relevant for that particular discipline (Losee, 1997; Shannon, 1993). Shannon states:

> *The word 'information' has been given different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of*

[1] Despite the fact that the idea of 'information' as a form of 'organization' seems to be a commonly used one, it is not accepted by everyone. For instance, Baudrillard claims that "INFORMATION = ENTROPY" (Baudrillard, 1994, p. 86). For Baudrillard, the informational organisation entails a neutralisation of that which it is information about, and with that Baudrillard considers it to be entropy (Baudrillard, 1994, p. 86).

this general field (Shannon, 1993, p. 180).

However, the existence of various views on 'information' is not necessarily problematic. Aamodt and Nygård argue that 'information' is a polymorphic concept, where its definition depends on its (theoretical) context (Aamodt & Nygård, 1995, p. 193). The challenge of this chapter is therefore, from these myriad views on information, to provide the reader with a convincing view of what information is and how we relate to it. For this, I have chosen to start this chapter by examining the often referred to 'Data-Information-Knowledge-pyramid' (hereafter: DIK-pyramid) (cf. Zins, 2007; Rowley, 2007) — below I will explain this choice. From the examination of the DIK-pyramid, I infer a particular view on information and explain how it relates to the perceiving human agent. Following this, I discuss how personal information can represent us. Next, I discuss how this personal information plays a role in interactions between people.

2.2 From data, information, and knowledge to signs

In this section, I present the first part of the framework that forms the backdrop of this study. This part of the framework concerns the concepts of information, and its relation to the world and human beings. In order to show the reader how I came to this particular framework and why I made certain choices, I will guide the reader through the steps that I made. For this, I start at my initial starting point: the DIK-pyramid. The DIK-pyramid is a fruitful starting point because it provides an account of information and of data, as well as an account of the relation between these two. In the following sections I take a closer look at the DIK-pyramid and build further on the view from there.

The DIK-pyramid entails a relational structure between the concepts 'data', 'information', and 'knowledge' (see figure 2.1). The main reasoning underlying the pyramid is that data gives rise to information, and information gives rise to knowledge (cf. Zins, 2007). This information hierarchy is widely used in information sciences, up until the point where it is taken for granted (Rowley, 2007, p. 163). However, the pyramid is not a fixed concept and we can find several variations on it. One of the more common variations is the addition of an extra layer above knowledge, like ' wisdom' (see e.g., Ackoff, 1989). However, the wisdom layer is not often discussed or used by authors, if acknowledged at all (Frické, 2009, p. 133). Because 'wisdom' and any superlatives to it are not often used, nor have any additional value for this study, I stick to the more general use of only 'data', 'information' and 'knowledge' as basic categories, and I assume 'knowledge' to encompass potential superlatives like 'wisdom'.

Moreover, the concept of the pyramid itself, as well as its hierarchy, can be understood in different ways and are a topic of discussion. One of the discussion points is the shape of the pyramid. Given the general view that information is inferred from data in a certain frame of interpretation, information adds more to

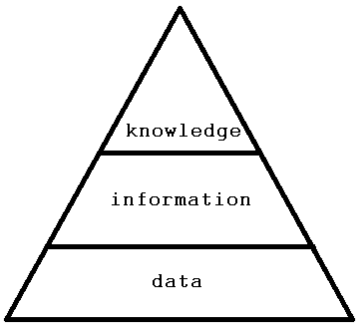


Figure 2.1: DIK-pyramid

the interpretation of the data than which can solely be inferred from the 'plain' data. This means that information is more extensive than data and also irreducible to data (Frické, 2009, p. 140). The same goes for the information-knowledge relation. This makes the DIK-pyramid as a shape rather misleading since the shape suggests a loss of mass when subtracting one layer out of the previous one; an upside-down pyramid would be more appropriate. Also, with regard to the different layers, the pyramid is not conclusive. The distinction between data and information as well as between information and knowledge often remains vague (Boisot & Canals, 2004, p. 44). Additionally, there are authors who argue in favor of a reversed hierarchy (cf. Tuomi, 1999), or who even want to abandon the DIK-pyramid altogether (cf. Frické, 2009).

Despite these uncertainties of, and disagreements about, the DIK-pyramid, the pyramid does have its merits by pointing out the different layers; while 'information' may be often used interchangeably with 'data' as well as with 'knowledge', 'knowledge' is not used interchangeably with 'data' (Boisot & Canals, 2004, p. 44). There is a general consensus that at least a step is required between data and knowledge and that knowledge cannot directly be inferred from data (Boisot & Canals, 2004, p. 44). Moreover, the three concepts seem to be all needed, and there is a certain relation between them.

Unfortunately, the DIK-pyramid as such does not bring us any closer to a definition of information. As with other literature, the literature on the DIK-pyramid is no exception when it comes to the many variations in definitions of 'data', 'information' and 'knowledge', as these tend to vary per scholar and discipline (cf. Frické, 2009). However, given the fact that the pyramid does seem the most plausible starting point to get more grip on the concepts 'information' and 'data' and the relation between these two, I will take a closer look at the definitions used as part of the DIK-pyramid. Though, given the issues that exist with the pyramid, I suggest focusing on the DIK-relations themselves and abandon the idea of the pyramidal shape.

2.2.1 Different views

As pointed out, the exact definitions of 'data', 'information' and 'knowledge' as well as the relation between them varies, even between authors who underwrite the DIK-pyramid (cf. Rowley, 2007). In order to get a grip on these definitions and their differences, I performed a meta-level analysis of the definition collection assembled by Zins in his article *Conceptual approaches for defining data, information, and knowledge* (Zins, 2007). For this article, Zins asked forty-four researchers on data, information and knowledge to provide him with their definitions of these three.

In the article, Zins concluded that the main difference between the many views on 'data', 'information' and 'knowledge' seems to be the realm where the phenomena exist: the subjective or the objective realm (Zins, 2007, p. 486). In the definition sets, we find at the one extreme the phenomenological approach according to which data, information and knowledge are phenomena that are fully dependent on a cognitive subject. At the other extreme, we find researchers who consider all three phenomena to exist objectively independent from any cognitive agent. To give some insight, I will give examples of the definitions of each phenomenon from an objective and a a subjective approach.

Data
(1) *Objective definition*: "**Data** are unprocessed, unrelated raw facts or artifacts" (Twining in Zins, 2007, p. 486).
(2) *Subjective definition*: "**Data** are sensory stimuli that we perceive through our senses" (Baruchson-Arbib in Zins, 2007, p. 480).

Information
(3) *Objective definition*: "**Information** is knowledge recorded on a spatio-temporal support(Le Coadic in Zins, 2007, p. 486).
(4) *Subjective definition*: "**Information** is the change determined in the cognitive heritage of an individual. Information always develops inside of a cognitive system, or a knowing subject. Signs that constitute the words by which a document or book has been made are not information. Information starts when signs are in connection with an interpreter" (Biagetti in Zins, 2007, p. 480).

Knowledge
(5) *Objective definition*: "**Knowledge** is the rules and organizing principles gleamed from data to aggregate it into information" (Hersh in Zins, 2007, p. 484).
(6) *Subjective definition*: "**Knowledge** is embodied in humans as the capacity to understand, explain and negotiate concepts, actions and intentions" (Albrechtsen in Zins, 2007, p. 480).

The different definitions lead to conflicting perspectives: e.g., according to (3) the content of a book would be information, while according to (4) it would not. The contradictions that arise between the different definitions are in many cases

the result of the ontological form attributed to data, information, and knowledge. This form determines their mutual relations as well as how and to what extent they can be perceived and transmitted. In order to get a stronger grip on the many different perspectives, I divided the definitions given in Zins' article into objective, subjective, and combined phenomena. By making this statistical list of the objective and subjective definitions, we can get some idea of what the more common views on the three phenomena are.

In order to create the list, I labelled every single definition (and thus not the complete definition set) with either 'objective phenomenon' or 'subjective phenomenon'. I considered everything that was defined as bound *within* a cognitive subject as a subjective phenomenon. In case of doubt, I gave the objective interpretation priority, because in the cases where a phenomenon was defined as subjective, its objective existence was often explicitly excluded (e.g., "Knowledge cannot be communicated by speech or any form of writing, but can only be hinted at" (Gladney in Zins, 2007, p. 483)), whereas vice versa this was not the case. The main reason to do this, was to deal with the allocation of views on knowledge to either the subjective or objective realm. Deciding whether 'knowledge' was taken to be a subjective or an objective phenomenon was especially challenging because knowledge was often defined as a set of rules (see e.g., Hersh in Zins, 2007, p. 486). This means that knowledge was always dependent to a certain extent on the framework created by human agents — however being 'created by' is something fundamentally different from 'only exist in'. Underlying the allocation of knowledge I therefore asked the question: "can knowledge according to this view be externalised and transmitted between agents?" — if the answer was "yes" I considered it to have (at least potentially) an objective existence.

In many cases the definition of 'information' was the most ambiguous with regard to its ontological character. I therefore had to derive its ontological status from the related 'data' and 'knowledge' definitions. In the cases where the definitions of 'data'/'information'/'knowledge' were immediately combined I tried to separate them as much as possible. In a few cases I needed the definition of the other phenomena by the same researcher to be able to make out whether the researcher saw the phenomenon as subjective or objective. Also, many definitions described the phenomena as a combination of subjective and objective aspects. For example: "Information, as a phenomena, represents both a process and a product; a cognitive/affective state, and the physical counterpart (product of) the cognitive/affective state. The counterpart could range from a scratch of a surface (...) [to a] written document" (Debons in Zins, 2007, p. 482). I labelled these combination definitions as attributing an objective as well as a subjective existence to the phenomenon in question. This explains why in all three cases the total percentage is over the hundred percent. Furthermore, some of the definition sets were incomplete.[2] In these cases, I only counted the definitions that were there. A last disclaimer: not all definitions were clear about the ontological status that they attributed to the phenomena, which may have resulted in an erroneous

[2]For example, Rousseau does not give a definition of knowledge (Rousseau in Zins, 2007, p. 486).

interpretation on my side. My labelling should therefore not be taken as a hard truth, but as a rough reflection on the differences in perspectives.

With the guidelines described above employed, I counted the following occurrences:

Phenomena	Objective phenomenon	Subjective phenomenon
Data	95 % (42)	20 % (9)
Information	80 % (35)	43 % (19)
Knowledge	36 % (16)	84 % (37)

The statistics show that data are the phenomena that are most often considered to have an objective existence, while knowledge is most often considered to be a subjective phenomenon that only exists in the human agent (this is also one of the conclusions of Zins himself (Zins, 2007)). Information is more often considered to have an objective existence than to be a fully subjective phenomenon. However, a note is in place here: in almost all definition sets, information as a phenomenon was heavily intertwined with either data or knowledge, and its ontological status generally depended on the ontological status of data or knowledge.

In order to figure out how to best understand data and information, especially in relation to the research at hand, I will examine the different perspectives in more detail. For this, I will first discuss the two extremes of a fully objective and a fully subjective perspective in the following subsections and point out their advantages and/or disadvantages.

2.2.1.1 The objective view

In the fully objective view data, information, and knowledge are regarded as completely objective existing phenomena. Of the roughly forty-four 'data-information-knowledge' definition-sets in Zins' article, seven were fully objective (Zins, 2007). An example of a objective view is the following definition set:

> ***Datum*** *is a quantifiable fact that can be repeatedly measured.* ***Information*** *is an organized collection of disparate datum.* ***Knowledge*** *is the summation of information into independent concepts and rules that can explain relationships or predict outcomes.*(Seaman in Zins, 2007, p. 486).

The ontologically objective view on the concepts of data, information and knowledge can be a useful perspective for some researches, like those that focus on the syntactic transmission of 'information'. For the analysis of a purely technological processing of information, it is therefore a usable view. However, the drawback of this view is that it does not account for the manner in which human agents perceive and interpret 'information' embedded in a specific context and how this interpretation may differ per agent given differences between agents in their backgrounds in for example, language, culture, education, etc. Because

the fully objective perspective does not take the human being as an interpreting agent into account, it is also not a very productive view for this study, because I seek to unravel the manner in which technology affects the relation between human agents and their understanding of the world.

Another difficulty of the more objective oriented definitions is that many of these frame data as 'facts' (see e.g., Floridi, 2016; Rowley, 2007). For practical purposes, I understand 'fact' here conform its definition in the Oxford English Dictionary[3]: "a thing that is indisputably the case. (...) the truth about events as opposed to interpretation"[4]. Defining data as 'facts' is problematic. Some authors point out that data in itself does not constitute facts, but that data already requires interpretation by being embedded in a theory or system or are the amplification of an earlier reasoning process (see e.g., Low, 2009). Data without interpretation in at least a certain language or system might not even be recognizable as 'data'. To give a contemporary example with regard to data in computers "[c]ode and data look the same in memory. They are only different in how you interpret them" (Duntemann, 1992, p. 113). Additionally, the presumed objective existence of facts as such is a challenge and might be more context and system dependent than we generally realise (Barad, 2007). Next to these difficulties, there is the trouble with fact-related definitions of data that faulty data — data which are not accurate — would not be data according to these definitions. This would mean that the recording of data can in retrospect turn out to be 'not data', because for instance faulty equipment was used. Taken that we often cannot know for certain whether data are correct, such a dependence on truthfulness in the definition is problematic (Frické, 2009, p. 137). This may provide a challenge in many contexts and is in my opinion the most important problem of the fact-definition of data. I have therefore chosen not to employ a view on data, information, and knowledge that has its roots in data as 'facts'.

2.2.1.2 The subjective view

In the fully subjective view, an agent's sensory and cognitive processes give shape to stimuli from the outside world in the form of data, information and knowledge. Two of the forty-four definition sets in Zins definition collection were fully subjective. The exact interpretation of the content of 'data', 'information' and 'knowledge' in this process varied per researcher. Here is one of the subjective definition sets as example:

> ***Data*** *are sensory stimuli that we perceive through our senses.* ***Information*** *is data that has been processed into a form that is meaningful to the recipient (...).* ***Knowledge*** *is what [has been] understood and evaluated by the knower*

[3]Other dictionaries give different definitions, see for example Merriam-Webster, which defines fact as "something that has actual existence", `https://www.merriam-webster.com/dictionary/fact`, last accessed 06-08-2019. Coming to a clear understanding of the word 'fact' and its use, is a topic of research in itself.

[4]Concise Oxford English Dictionary, Eleventh Edition, Oxford University Press.

(Baruchson-Arbib in Zins, 2007, p. 485).

The interpretation of the relation between data, information and knowledge in the above cited set of definitions are rather linear: data are processed into something meaningful to the agent and thereby becomes information, which in turn can lead the agent to gain knowledge. However, such linearity does not explain how an agent can actually infer information from data, and knowledge from information. For this, the model presented by Boisot and Canals may help us to get a more complete picture.[5] For the sake of clarity, I have reproduced Boisot and Canals' model in figure 2.2.

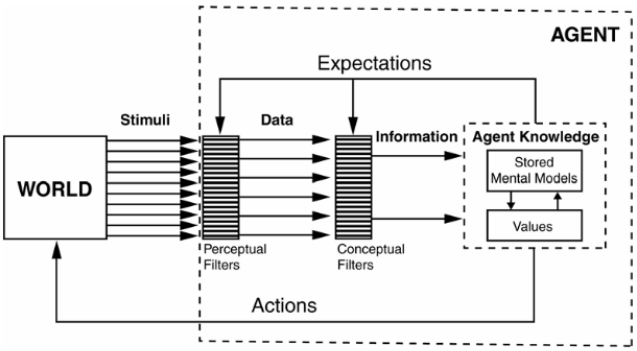


Figure 2.2: Figure from Boisot and Canals' *Data, information and knowledge: have we got it right?*(Boisot & Canals, 2004, p. 48)

According to this model, an agent receives stimuli from the outside world: everything in the world can be a stimulus for an agent. These stimuli enter the agent's cognition through her sensory perception (the agent's 'perceptual filter') and are registered as data (Boisot & Canals, 2004, p. 47). Data are thus stimuli that are consciously and subconsciously discerned by an agent, either without aids or with the use of technology (Boisot & Canals, 2004, p. 52). This registered data in turn is passed through a 'conceptual filter' which allows the extraction of information from the data (Boisot & Canals, 2004, p. 47). In order to infer information from data, the data needs to be interpreted. This interpretation "involves an assignment of the data to existing categories according to some set of pre-established schemas or models that shape expectations" (Boisot & Canals, 2004, p. 55). The information thus depends on the agent's frame of reference. As such, information has a relational character (Boisot & Canals, 2004, p. 52).

[5]It is debatable whether Boisot and Canals hold a fully subjective view. The scheme used in their article suggest a fully subjective view, while their text at certain points suggest data to have an objective existence *in the world* (compare Boisot & Canals, 2004, p. 48 and p. 63). Because I found their scheme valuable for clarification purposes, I will use their view as being fully subjective — although duly noted that they may have meant it otherwise.

The agent's frame of reference is constituted by the agent's "cultural background, unconscious intuitions, concrete memories of similar observations in the past, expectations triggered by the specific context, as well as text book knowledge and domain dependent heuristic rules" (Aamodt & Nygård, 1995, p. 198) — ergo, the agent's 'knowledge'. The process of registering data and extracting information is thus highly shaped by the agent's prior knowledge, which consists of the agent's stored mental models and values (Boisot & Canals, 2004, p. 48). An agent's knowledge is "an inherent resource of a reasoning agent that enables inferring of new information from existing information [that] (...) may, in turn lead to the inferring of more information, and so on" (Aamodt & Nygård, 1995, p. 199).

Agents will differ in their knowledge, because "agents always will have different histories, experiences, environments of operations, etc." (Aamodt & Nygård, 1995, p. 202). Taking into account that the meaning given to information is shaped by the context in which it is interpreted, and this context varies across individuals, individuals thus infer different information from the same set of stimuli (Boisot & Canals, 2004, p. 53). Yet despite the subjective character of knowledge, a form of common knowledge can still exist if agents have a similar background (cultural, theoretical, contextual, etc.) which equips them with roughly the same frame of reference (Aamodt & Nygård, 1995, p. 202).

The fully subjective view on data, information, and knowledge, is compelling. Unfortunately, it leaves the focus point of this study hanging mid air. Defining 'data', 'information' and 'knowledge' within only the cognitive faculties of an agent does not indicate the manner in which 'information' is made tangible for our perception and transmittable through objects and the like. Reducing all the 'input' to undifferentiated stimuli leaves us with little ground to evaluate the manner in which these stimuli are constituted, shaped and posed for an agent's perception by things in the world. As such, the fully subjective view is of little help in being able to expose how our interpretation of stimuli may be affected by the technologies that bring them forth — and therefore is on itself too narrow to work with for the purposes of this study. However, by combining it with objective elements, it can become a viable theory to assess the issues that lie at the heart of this study.

2.2.1.3 Towards a combined view

The challenge with the purely subjective and objective definition sets, is that they are embedded in a particular bilateral relation, i.e. in the relation between information and human cognition, or in the relation between the world and information. However, these models do not need to rule each other out, and can even coexist and complement each other. Models that combine these two relations tend to bridge the bilateral relations by accounting for an objective and a subjective side of data and information. As such, they account for the manner in which human beings acquire data, information and knowledge in their cognitive processes, while they also account for the manner in which information is transmitted and presented 'in the world' between agents.

Most definition sets in Zins' article consist of a combined view with regard to

the ontological status of data, information, and knowledge. These views vary to a greater or lesser degree from each other. In some views, data, information, and/or knowledge are all three considered to have both an objective as well as a subjective existence. However, many definition sets take data to be objectively existing phenomena and knowledge to be a subjective phenomenon. Information hovers in between as either an objective phenomenon, a subjective phenomenon or a combination of both. Multiple definition sets present the idea that knowledge itself is not transferable between agents, but that information allows us to communicate about our knowledge to each other (see e.g., Wersig & Neveling, Gladney, and le Coadic in Zins, 2007). This in turn requires information to be — at least till a certain extent — an objective and transmittable phenomenon.

The merit of the combined views is that they generally account for the role of the human being as interpreting agent, while equally providing grounds for looking into the manner in which 'information' is processed outside of the cognitive agent. This is beneficial, because this allows us to research what happens with information and data outside of the human agent. Frické states: "the core of information science is still the attention to external storage, storage outside the 'skinbag' (...), that is, to those artefacts of preservation that form the bridge from the individual and instant of time to availability across individuals and persistence through time" (Frické, 2009, p. 138). The same counts for this study. In the next section, I therefore present an account of data and information from a combined subjective/objective view. The presented model will serve as an analytical instrument for the research in the upcoming chapters.

2.2.2 Data, information, knowledge and the human agent: a model

In this section, I explain the view on data and information that I will use. As discussed in the previous section, this is a 'combined view'. However, coming to a model that can be used as part of a toolkit to analyse the manner in which information technologies can affect our interactions with and perception of personal information, is not a case of just adopting a set of useful and plausible definitions. Instead, it requires a theory that accounts for several aspects and is usable to further explore the issues at hand. In order to construct such a complete picture, I found it necessary to combine several theories. I tie in to the subjective view of Boisot and Canals, but will elaborate and complement it with some more objective components to account for the differences in information transmission by means of technologies, which will become important later in this study. Moreover, the differentiation between data and information that I employ is a practical choice, but not necessarily the only plausible one.

My starting point is that, at the very least, the outside world is a source of data and information. I follow Boisot and Canals, amongst others, in their standpoint that data "and the regularities that reside within the data, are properties of events and things 'out there' in the world (...) that become available to us as sentient beings through our physiological apparatus, often amplified by instruments and

other artefacts" (Boisot & Canals, 2004, p. 52). As such, data are anchored 'in the world' and thus have a certain 'objective' existence. A datum is best seen as "the smallest collectible unit associated with a phenomenon" (Haythornthwaite in Zins, 2007, p. 483).

In turn, information is an "assessment or interpretation of data" (Haythornthwaite in Zins, 2007, p. 483). It is an abstraction of data, that "does not inherently mean empirical or first hand analysis of data. It also does not guarantee correct interpretation of data although that is expected" (Haythornthwaite in Zins, 2007, p. 483).[6] While interpretation is dependent on an interpreting agent, I argue that the potential of information is *also* already out there in the world. For this, I make a small sidestep to Gibson's concept of 'affordance' (Gibson, 2014). This concept is productively used by several authors to discuss the relation between humans and technology (see e.g., van den Berg & Leenes, 2013; Hildebrandt, 2015). However, I think this concept also is helpful to explain the relation between humans and their surrounding world in terms of information.

Originally Gibson's concept of 'affordance' is aimed at the relation between subjects and their environment taken from an environmental psychologists perspective. The agent's environment 'affords' certain things to the agent, it is what the environment "*offers* the animal, what it *provides* or *furnishes*, either good or ill" (Gibson, 2014, p. 119). The affordances are constituted by the properties of (parts of) the environment of an agent (Gibson, 2014, p. 119). There seems to be no limit to what can constitute an affordance; environmental elements like surfaces, as well as objects or other agents can constitute affordances for an agent (Calo, 2016, p. 5). An agent can perceive what the agent's environment affords (Gibson, 2014, p. 112). These affordances are measured *relative* to the agent, and are therefore *unique for every agent* (Gibson, 2014, p. 120). They are dependent on the action capabilities of the agent (McGrenere & Ho, 2000, p. 181). This implies a strong subjective component to the concept and it is one of the key insights of affordance theory (Calo, 2016, p. 2). However, next to the subjective component, the concept of affordance also sees to a part which is anchored in the physical world and is thus also in a sense objective (Gibson, 2014, 121). As such, the environmental affordances "cut across the subjective/objective barrier. They are objective in that their existence does not depend on value, meaning, or interpretation. Yet they are subjective in that an actor is needed as a frame of reference" (McGrenere & Ho, 2000, p. 180). An affordance is a reciprocal interplay between the world and an agent's *perception*: "what the environment 'provides' in the way of 'stimulus' is a function, in part, of the organism's activity" (Sanders, 1993, p. 288). An agent may or may not perceive certain affordances of an object, but that does not change the affordances of that object. Affordances are "properties of things *taken with reference to an observer*" (Gibson, 2014, p. 129). As such, an affordance is a non-neutral possibility that with its existence,

[6] I do not claim that these definitions allow for a strict demarcation between 'data', 'information', and that is not their function. The concepts are tools that allow us to analyse the manner in which online 'personal information' can affect our knowledge of each other and ourselves.

once the agent is aware of it, co-shapes the agent's world-view by providing the agent with certain options.

If we now look at data and information in this context, we can say that our environment affords us to infer data and information. The manner in which an agent perceives the affordances of certain data, will shape the base for the agent's use of that data. Moreover, whether something is data or information, depends in part on the perceiving agent. For instance, a piece of paper with text has certain informational affordances. However, whether we are able to infer a certain message from the paper depends on whether we can recognise the representation and understand the language. If not, we can only infer data. Data and information are thus anchored at least partially in the world outside the agent — with the disclaimer that they will always need to be interpreted as stimuli by an interpreting agent, and therefore are inherently intertwined with an agent's subjective interpretation.

Lastly, I tie in to a significant part of the DIK scholars who argue that knowledge is fully subjective. In this, knowledge is:

> *what an individual takes from information and data, and what they incorporate into their beliefs, values, procedures, actions, etc. It is heavily internally oriented, understood completely only to the person possessing it. Much work around knowledge implies how to get the knowledge "out of" one head and in to another. Such transfer entails encoding knowledge into transferable information and decoding again into knowledge. Knowledge and information are not the same, but they feed from and support each other* (Haythornthwaite in Zins, 2007, p. 483-484).

With this view on the concepts of data, information, and knowledge, in the back of our mind, we can now delve more deeply into the relation between the external anchoring of information in objects, such as a postcard, and the human interpretation thereof.

2.2.3 Signifying object and subject

I start this subsection by examining how data and information are anchored in the outside world. Next, I add the human interpreter to the mix and combine this in a model loosely based on Peirce's semiotics.

Wiener stated: "Information is information, not matter or energy" (Wiener, 1961, p. 132). Despite this, information does often relate to matter and/or energy by means of being 'carried' by it; we send messages on paper, over the wire or through sound. By means of these carriers, we can perceive data and information by means of stimuli. The carriers are the prime focus of this study. However, because this study concerns online objects that provide users with personal information and not a bunch of separate data (which could not yet be recognised as relating to an individual), I focus on carriers of information, and not the carriers of merely data. I will explain this with an example. Take for instance a digital image showing a person. The combination of the data into the information

'person' is an information affordance of the object. Data in this case could be for instance a black pixel at location (383, 937), a white pixel at (740, 217), etc. However, in the object the pixels are already combined in a bigger picture that allows us to directly infer information about a person. E.g., we could infer that it is a man, we identify some traits of his appearance, and maybe we can estimate his age. The object thus provides more information than only the location and colour information of separate pixels. These separate data (in this case pixels) combined on a particular carrier, thus give rise to an image that represents information about the man's appearance.[7]

Because I am interested in objects like photographs and news articles, and not in separate pixels, I take these objects to provide users with information. Only in exceptional cases, these objects would tell a user nothing and not transcend their status of data. However, this would mean, in case of for example an article, that a user would not even be able to recognise that the object represents an article or a language. These cases seem so rare, that for the purposes of this study I can assume that online objects provide us with information, even if it is limited to the recognition that it is something conveyed by a language that we do not understand. In this context, the relation that I research at this stage is the relation between the perceiver (the 'user' of the information), the information, and the thing or person to which the information refers.

In order to get a better understanding of what goes on in this relation, especially with regard to the transmission and representation of information by objects, I let myself be inspired and guided by the semiotic theory of Peirce to explain what I take to be the information-conveying structure of such objects. However, I will not follow exactly in Peirce's footsteps, and even make some radical alterations to his terminology. The reason for this is that, first of all, Peirce's theory is too extensive and too complex to discuss here in depth. Secondly, because Peirce shifted over time in the use of his terminology (see e.g. Jappy, 2013), I decided for clarity purposes to utilise one set of my own working terms that can work with the various other theories in this study. I will not enter the discussion or compare the (different) terms of Peirce and their uses. Thirdly, I add the 'material' dimension of information carriers to the mix, thereby bridging Peirce's semiotic model and the philosophy of technology. I will discuss this further in chapter 3. Nevertheless, I have chosen to use Peirce's model — albeit in a simplified and adapted version. I made this choice, because Peirce's model transcends linear views on signs by incorporating the element of the subjective interpreter (Chapman *et al.*, 2004, p. 385). This allows me to explore the manner in which an object, like a drawing on a piece of paper, relates on the one hand to what it represents, and on the other hand to the human agent who interprets the object.

[7]Because the human perception is the point of departure, I employed this particular division of a picture in data and information. However, I can image that in another context, the identification of data and information would be different. For example, in a large database with satellite photographs, a single photograph may be considered a single datum. On the other hand, when working on a microlevel with pictures, the RGB value of a pixel may be a datum, while a pixel itself is already seen as information.

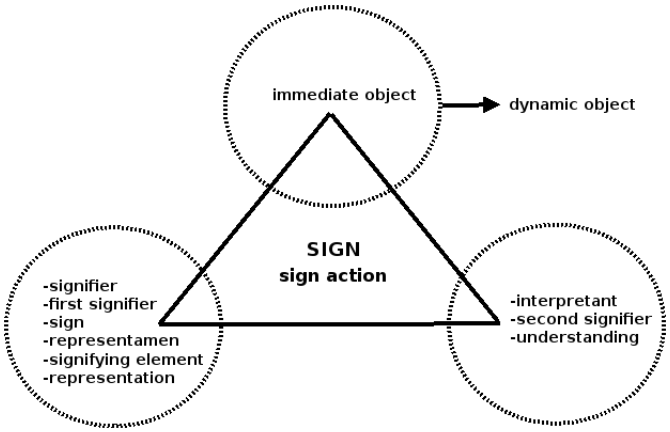


Figure 2.3: Terminology of Peirce, as well as some of the secondary literature

Peirce's semiotic theory is grounded in a triadic relation that constitutes a 'sign' (Peirce, 1974, CP 2.242). Peirce understands the concept of a 'sign' as an 'action' being comprised by three core elements that are related in a three-way to each other; 1) a signifier, also called 'representamen' and 'sign' by Peirce (please note that Peirce uses the term 'sign' double; he uses it for the signifier as well as for the complete sign), 2) the object that it represents and 3) the interpretant (Peirce, 1974, CP 1.480, CP 5.484). Figure 2.3 shows a rough view of Peirce's triad and the various terms used with regard to the three elements. I will now discuss these three elements of the triad and my adaptations thereof.

Signifying object The signifier, the thing that it represents, and the interpretant are highly intertwined elements that shape a sign together. The focus of this study lies on a particular element: the signifier. The signifier *mediates* between that to which it refers and the impression that it makes in the interpreting agent (Peirce, 1998, p. 276). It is that which represents something for an agent (Burch, 2018). I argue that to be perceivable for an interpreting agent, the signifier necessarily has a certain material existence.[8] I understand 'material' here in the broadest sense, and take it to also cover things like vibrating air in the form of sound. This material form affects the signifier's relation to the perceiving agent and the thing that the signifier represents. For example, a signifier referring to a person differs fundamentally if it is in the form of text or a photograph; it differs in how and when we can identify the person by means of the signifier. Text may allow us to identify a person by her name, or her phone number, while a photo allows us to identify a person based on her appearance. While the object carrying the signifier and the information it affords are not exactly the same thing, they do enter our perception as one set of stimuli. The materialisation thus sets the base for the

[8]I will leave hallucinations outside the scope of this study.

form of the information and shapes the stimuli that the agent perceives. With this, the materialisation affects the perception and interpretation of the information. It is the materialised signifier that is interpreted by the agent who perceives a sign, and not the entity to which it refers, nor the intention of the agent expressing or creating the signifier. I will discuss the impact of the materialisation of information further in the next chapter.

The thing that is represented by the signifier, as well as its interpretant, are thus necessarily tied to the perceived existence of the signifier: without a signifier that enters our cognition by means of stimuli, there is no reference, nor something that gives rise to an interpretant. With a foresight of what is to come in this study, I will call the signifier in its material form a *signifying object*. This is the point where I break with Peirce. For the purpose of this study, I found it necessary to reassign the term 'object' to another part of the triad, and change his 'object' element to 'subject' (which I will discuss in the next paragraph). This switch allows me to streamline the terminology with the GDPR, as well as with other theories, and thereby establish a readable and consistent terminology throughout this study.

Lastly, as the reference and the interpretant both depend on the existence of the signifying object, and the signifying object fails to *signify* without a reference and interpretant, I will use 'signifying object' for the signifier as well as for the total of the sign (this is similar to Peirce, who uses 'sign' for the total triad, as well as for the signifier).

Reference and referent or subject In order to signify something, a sign relates to something outside the sign; this can be anything from a concrete human being to a feeling, an idea or a fictional event (Nöth, 2011, p. 29). As such, even self-referential signs relate to something outside the sign, namely an abstract idea. To explain how a signifying object signifies, Peirce distinguishes between that which is represented with the sign, this is what Peirce calls the 'immediate object', and the thing as it is outside the sign, which is roughly what Peirce's calls 'dynamic object' (Jappy, 2013, p. 14). I will follow this distinction, albeit I will employ a simplified and somewhat adjusted version.

Let me start with that to which a sign refers, the thing outside the sign that Peirce calls 'dynamic object'. I will call this the *referent* and *subject* interchangeably, or just simply use the noun that is used for the entity outside the sign, like 'the individual'. I chose to radically diverge here from Peirce and use the term 'subject' to pave the way for the discussion of art. 17 GDPR. The GDPR uses the term 'data subject' for the individual to whom information refers. As such, using the term 'subject' streamlines the terminology in this study. Moreover, the use of the word 'subject' underlines the role of personal information about an individual. As a subject of information, the individual is in a sense 'subjected' to the sign action: in the the eye of the interpreter of the signifying object, the individual is subjected to the information afforded by the sign. This information constructs the interpreter's understanding of the subject. However, an interpreter will never have all the information about a particular subject: even if she knows all

the information that is available now, there is always the chance of new information coming to light in the future (Nöth, 2011, p. 30). The referent therefore "belongs to a reality independent of its sign to which we have no full access" (Nöth, 2011, p. 30). Lastly, it is important to note that the referent does not have a strictly material connotation and can be something immaterial or fictional, like a theory, a feeling, a theatre play, an idea, or a fictional entity like a unicorn (Nöth, 2011, p. 31).

While the referent itself is not part of the sign, the sign does refer to the referent in a certain manner. This brings me to Peirce's notion of the 'immediate object'. This entails "the Object *as* represented in the sign" [my emphasis](Peirce, 1998, CP 8.314). The immediate object is connected to both the signifying object and the interpreting agent: it is what we know about the referent from the sign (Nöth, 2011, p. 30). However, this information is always partial as the real referent can never be fully represented in a sign and can even be erroneous or falsely represented (Nöth, 2011, p. 30). For the purposes of this study, I will call the object as represented in a sign, a *reference*[9]. I connect here to the work of Ihde, who describes 'reference' as that which is hermeneutically made present by a technological object (Ihde, 1990, p. 91). I will discuss the manner in which technology presents references to us further in the next chapter.

It is important to note that with this terminology I make a (possibly somewhat counterintuitive) distinction between the reference to a particular referent *as* represented in the signifying object, and the concrete representation of the referent by the signifying object (compare the two different points of the sign triad that Peirce attributes to 'representation' and 'immediate object', see figure 2.3). I take the reference to be that what a signifying object signifies about a referent, i.e. the information it affords, while the representation by the signifying object is the material form in which this information is given shape. For example, two different representations of one referent are a photo of my cat and a text in which I describe how my cat looks. However, because these two representations both afford an interpreter roughly the same information about the referent, 'the cat is black with green eyes and pointy teeth', I argue that they share a certain reference (a particular appearance of the cat), but in another material form. While the reference to and the representation of a referent have a significant overlap, distinguishing between these two dimensions allows me to better delve into the informational impact of certain signifying objects later on in this study.

Interpretant The last element of the sign is the interpretant. The interpretant is "a mental concept produced both by the sign and by the user's experience of the object" (Fiske, 2010, p. 42). It is the understanding that the human agent has

[9]Peirce seldom uses the term 'reference' in this theory. When he uses it, he uses it to point out the three relations that take place within the sign: object - interpretant, sign - object, sign - interpretant (Peirce, 1998, CP 5.283). For an interesting paper on the relation between the terms 'reference', 'referent' and the semiotic work of Peirce, I would like to refer the reader to *Representation and Reference According to Peirce* by Nöth (2011). While I use the term 'reference' not conform Peirce, I choose to use it instead of 'immediate object' for the clarity of the overall text in this study.

of the relation between the sign and the referent (Burch, 2018). The interpretant is based on ideas that already exist in an agent's memory and which are being addressed by the sign (Short, 2007, p. 30). The interpretant thus depends on the interpreter. The interpreting agent therefore plays an important role in the constitution of what a sign represents (Chapman *et al.*, 2004, p. 385). Because no human being has the exact same experience as another, different interpreters will always interpret a sign to some extent differently (Jappy, 2013, p. 7).

The interpretant is a necessary element of a sign, because the sign only signifies something by being interpreted as such (Short, 2007, p. 30). This ties in with the earlier discussed theory of affordances: the sign is in its signifying meaning always dependent on a particular representation in the world, as well as on the interpretation thereof by an agent.

Information about a kiwi In short: the signifying object refers to a particular referent by means of a reference. This reference needs to be interpreted by the agent that perceives the signifying object. I have displayed the different elements and their mutual relation in figure 2.4. Because all three elements are necessarily a part of the sign, I placed them within the triad. The three elements that constitute a sign relate to the model of information presented in the previous section. First of all, in both perspectives the interpreting agent plays an important role in the meaning that is given to the information. Secondly, the information is a representation of something that exists outside the information. In this, I take information to be the part of the sign that exists in the outside world as an object that gives off a particular set of stimuli, and which in turn affords an internal counterpart where the stimuli are interpreted.

In order to clarify how I will use the terms of the model throughout this study, I will explain my adaptation of Peirce's model for a sign action that goes awry. In figure 2.4, we can see a postcard with the message: "I like kiwis!". The sender of the postcard is what I will refer to as the *expresser*. The postcard is a signifying object. The signifying object contains the reference 'kiwi' in the form of text. The reference is thus an inseparable part of the sign. The expresser who used the reference 'kiwi' intended to refer to the bird. As such, the real life kiwi bird is the intended referent outside the sign. However, whether an individual understands the word 'kiwi' as a fruit or a bird, depends on the specifics of the representation, the context, as well as the experience and knowledge background of the interpreting individual. The mental image that the reference as presented by the signifying object produces in the interpreting agent, is the interpretant. In this case the receiver of the postcard has a different association with the word 'kiwi' than the sender. As such, she interprets the signifying object as something that refers to a love for a certain kind of fruit. The result of this particular sign action is that the representation of the reference leads the receiver to form a view of another referent than the expresser intended to express with the signifying object.

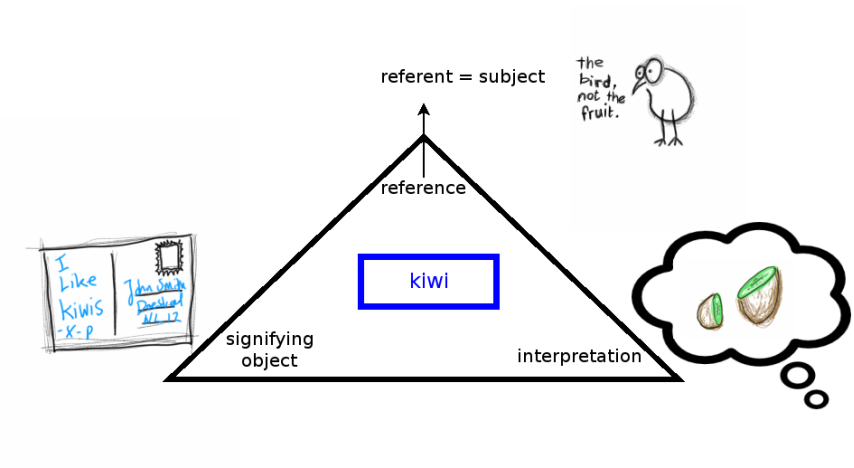


Figure 2.4: Signifying object, subject and interpretation

2.2.4 The presence of information

Information thus exists — at least partially — in the world outside us. By means of signifying objects, it can "make present what is absent in time and/or space" (Hildebrandt, 2015, p. 38). Such objects can refer to people, things and events that are often physically not around, lie in the past, do not even exist anymore, or may have never existed. While these objects do not make the literal subject present, they do imbue the referent with a particular 'presence' in the form of a reference. I define the 'presence' of information here as the *the afforded quantitative and qualitative proximity of a specific reference in time and space for human agents compared to other references*. Let me clarify this with some examples. Let us say that I have a picture on my desk portraying my cat and my partner together. This picture is a signifying object with a reference to my cat as well as to my partner. While they are both not near me at the moment, the particular references to them as represented by the picture are present for me. So far, the presence of the signifying object and the references coincide. However, let us now assume that I am a crazy cat lady. Next to the one picture of my cat and partner, I have nineteen other signifying objects on my desk representing my cat. These are miscellaneous objects: drawings, photographs, a poem, and even a sculpture, all referring to my cat. We now have twenty signifying objects, but still with only two references: a reference to my partner and a reference to my cat. However, there is a vital difference in their presence. While my partner is represented in one signifying object, the cat is represented in twenty. The reference to my cat is thus more often made present for my perception than the reference to my partner; it has a higher quantitative proximity. I therefore argue that the reference to my cat in this context has a *higher presence*. Now let us assume that I at a certain point decide to place one extra photo of my partner on my desk, but instead of the relatively sober

framing that I used for the twenty other objects, I enlarge this photo and frame it in a highly attention attracting frame with flickering lights. While still only two signifying objects refer to my partner, one of these objects clearly stands out and signals importance due to its framing. This object is made prominently present and thereby imbues its reference with a certain qualitative proximity. As such, the reference to my partner gains a high presence, possibly even exceeding the presence of the reference to my cat. In this manner, the presence of a particular reference for us is constructed by the presence of signifying objects in diverse quantities and forms; as the carriers of the information, they affect how and when we encounter particular references.

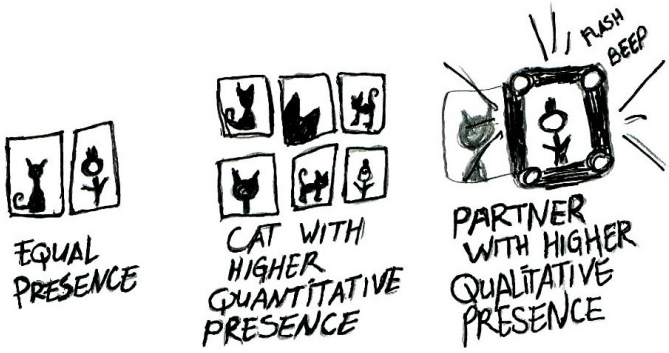


Figure 2.5: Presence of a reference

Moreover, the constitution of the presence of a particular reference is not only dependent on the signifying object that carries it, but also on the manner in which the signifying object is embedded in the bigger informational environment, which can be a house, a library, the Web, etc. As such, information may be more or less difficult to access, or may stand out or not. The effort that agents need to spend to access the information depends on the friction caused by the characteristics of the signifying objects and this environment (Floridi, 2005, p. 186). To give an example, online information is relatively easily accessible when you have a device with an internet connection, but more difficult to access when you do not have such a device, because you then first need to get access to a device with which you can access the Web. The presence of information is therefore also dependent on the resources, skills and capacity of an agent.

Given the role of the experiencing agent, I hold that the presence of information is rooted in what Ross calls 'existential space' (Ross, 2013). Ross explains 'existential space' as follows:

How near and far things feel is not merely a matter of distance. I can walk a kilometre very easily, but not if it is up a mountain, or through five feet of snow, or if I have a bad leg. Similarly, the places which are most familiar to me—my

home, my street, my office—are not merely objective geometric spaces, they are familiar regions marinated with memory and meaning. Familiarity with places is what makes them ready-to-hand, it is why they feel intimate, comfortable, and 'homely' (Ross, 2013).

The presence of information is thus strongly dependent on the perceiving agent. Her embodiment as well as her background shapes her potential interaction with the world: "Our everyday spatial involvement implies a pre-thematic sense of where things are, where we are in relation to them and how accessible they are" (Ross, 2013). The manner in which a reference is made present by signifying objects, affects the chance that a specific agent is exposed to it and that it grabs her attention. Moreover, it affects her understanding and interpretation of the reference — and thereby of the referent. For example, if a signifying object is prominently placed (e.g. an article on the front page of a newspaper), it signals importance and people are likely to interpret the information it reveals as something that is considered to be valuable or relevant to know.

If a certain reference is strongly present in the world, the chances are increased that it grabs an agent's attention. However, it is important to note, that even a strongly present reference does not necessarily grab an agent's attention in an equally strong manner — if at all; the agent may be unable to understand or recognise the reference, or may actively choose not to pay attention to it (note that this does not mean that it will have no conscious or unconscious effect on the agent at all).

2.3 Personal information and the informational persona

Appreciating the power of information to analyse people as well as to predict and even control their actions is not new; it is the very essence of human social relations and interaction.

Helen Nissenbaum, *Privacy in Context*, 2010

So far, I have formulated a perspective that allows me to examine the relation between information in the outside world in the form of signifying objects, the subject to which they refer and the interpreting agent. Now, it is time to zoom-in on the particular type of information that lies at the heart of this study: personal information. In this section, I discuss what I mean with 'personal information', and discuss why the manner in which personal information is conveyed by objects matters.

2.3.1 Personal information

Signifying objects can hold information about a particular person, and thereby reveal 'personal information' to an interpreting agent. For the purposes of this study, I align the concept of 'personal information' used here with the GDPR. However, I first need to point out that there is a difference between the GDPR and the concept employed here: while the GDPR speaks of 'personal data', I have instead chosen to focus on 'personal information', as explained in section 2.2. However, my approach is not necessarily contrary to the GDPR, because the GDPR defines 'personal data' as "any *information* relating to an identified or identifiable natural person" [my emphasis] (art. 4(1) GDPR). It lies outside the scope of this study to examine and clarify the concepts of 'data' and 'information' and their mutual relation in the GDPR. I will therefore stick to the concept of information, unless I quote the GDPR which seems to have put 'information' under the umbrella of 'data', or if I really mean data in the sense of separate datums that need to be combined and interpreted in order to form information. I hope it will be clear for readers from the context which of the two uses of 'data' are the case.

'Personal information' is information that refers to a particular person. Following art. 4(1) of the GDPR, it entails:

> *any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*

An important element of personal information, is that the information relates to a particular identifiable individual. 'Identifiable' in this context does not mean that an individual *is* identified, but that it is *possible* to identify her.[10] Such identification requires a description or a sign that sets one person aside from others and allows us to trace this back to a particular human being.

Many of the signifying objects in the world contain information relating to an identifiable individual. Think of photographs with recognizable people, newspaper articles in which specific people are mentioned, documents about school attendance, names signed on a petition, etc. All the references embedded in these signifying objects that refer to a specific individual form together what I will call an individual's *informational persona*.[11] The informational persona consists of references in all signifying objects, varying from a handwritten letter in a shoe box in someone's attic, to a digital photograph published on an online newspaper's

[10] WP 29, Opinion 4/2007 on the concept of personal data; p. 12.

[11] This concept has been inspired by inter alia Clarke's and Roosendaal's discussion of the 'digital persona' (Clarke, 1994; Roosendaal, 2009). However, because the 'digital persona' missed some aspects which I need in this study to employ the persona as an analytical tool, I decided to use a slightly different concept.

website. The presence of the references shapes the appearance of the informational persona for the experience of agents. Highly present objects will thus have a stronger impact on the appearance of the informational persona than less present references — albeit the exact presence for a particular agent will depend on the agent's own background, capacities and actions.

It is important to note that the informational persona consists of *all* the references in the world that refer to the individual. The informational persona therefore does not necessarily comprise of a one-on-one connection to the referent's identity; references may wrongly reflect the real life person of the referent, or may fail to show certain parts that the individual herself believes are vital to her identity. Outdated, inaccurate or even incorrect references are also part of an individual's informational persona as long as they refer to the individual in an identifiable manner.

2.3.2 The impact of the informational persona

The presence of references to a particular person can heavily affect an individual's identity construction in the eyes of others as well as the experience of the self. The signifying objects allow observers to take account of particular references, interpret them and infer predicates that construe a view of the individual in their perception. The predicates can consist of anything, ranging from factual elements like name and birthplace to subjective judgements, characterisations and classifications like 'funny', 'mean', 'stupid', 'criminal', 'hero', 'technoviking', 'father', etc. Signifying objects thus allow people to get a particular view on the informational persona, to interpret this and to attribute certain characteristics to a particular person. As such, the subject of the references is constructed as a subject in the eyes of the beholder. In turn, the responses of others based on these objects, can affect the self-perception of the referent. Moreover, signifying objects can also directly affect the self-perception of the referent herself by reminding her of past events, or confronting her with a view on herself that she did not see before. In this section, I will explain the relation between signifying objects, the informational persona, the referent, her identity, and others in more detail.

The agent's experience of an individual's informational persona (which can be her own informational persona), and her corresponding understanding of the subject, will depend on the signifying objects that the agent comes across, as well as their form, and the context in which she experiences them. It is impossible that any agent will perceive an individual's complete informational persona, if only for the simple fact that there will be no agent that has access to all the references referring to a particular referent in all the objects in private collections, as well as in all public, corporate and governmental collections. Moreover, the presence of the diverse references will depend on the perceiving agent's background and context. The result is that an agent will always have a certain perspectival view of an individual's informational persona (see figure 2.6).

This outlook of an agent on the informational persona is extremely important, because information plays a fundamental role in human interaction. Since we

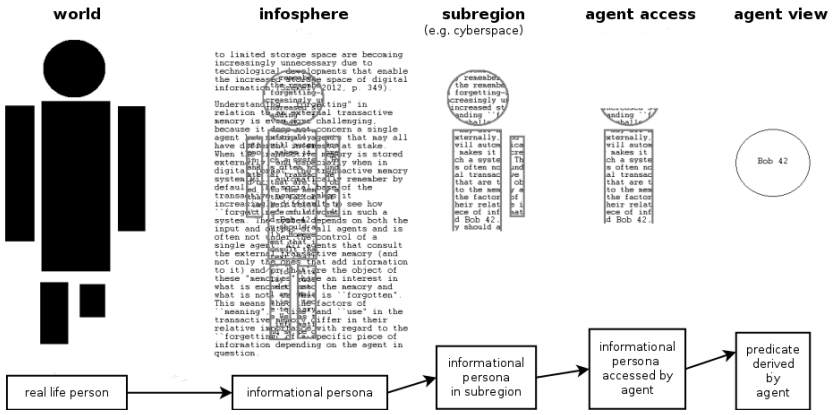


Figure 2.6: The informational persona

cannot look in each others' minds, we are dependent on information we receive — either provided to us by people themselves or by other sources — to get an idea of the character and identity of the people that we are dealing with. People therefore respond to each other based on their evaluation of this information; they use the information to ascribe certain predicates to individuals like social attributes and categories and allocate a 'social identity' to her (Goffman, 1963, p. 12). In this, the cultural and knowledge background of the agent also plays a pivotal role; the agent's conscious and unconscious beliefs colours her interpretation of signifying objects, and thereby of the subject. The understanding that we have of each other, necessarily takes shape within our own frame of reference (Susser, 2016, p. 6). The same set of information (even plain factual information) can therefore give rise to various assumptions about an individual (Susser, 2016, p. 3). For example, the classification of an individual as 'woman', 'gender neutral', or 'man', leads to different assumptions about the referent by conservative and progressive agents. This impact of the agent's frame of reference on the manner in which she interprets the identity of a referent expresses a certain exercise of power as it "socializes, invites, and reproduces social distinctions that mark social prejudice" (Thalos, 2010, p. 81). As the perceiving agent tends to fill in the referent's social identity based on her own background knowledge and assumptions, the freedom that individuals have in the eyes of another to have a certain identity can be very limited (Sen, 2007, p. 28). Meanwhile, the (conscious or unconscious) interpretation of an individual's persona shapes the perceiving agents' actions towards an individual (Goffman, 1959, p. 21-22). The social identity that agents believe that an individual has, forms the ground for their normative expectations and demands of the individual (Goffman, 1963, p. 12). In turn, the manner in which these agents respond to the individual, is likely to affect the referent's self-perception (Falk & Miller, 1998).

When an agent attributes a particular social identity to the referent based on

an encounter with particular signifying objects, this is often a fraction of the full identity of the individual, because people generally do not have a singular identity, but often have multiple identities that show their belonging to various groups (Sen, 2007, p. 26). For instance, depending on what online objects referring to me people encounter, they could see me as a philosopher, a legal scholar, a vegetarian, a Dutch, a World of Warcraft player, a female, a caucasian, a punk, a nerd, an opera lover, and so forth. Unfortunately, once an agent attributes a particular identity to an individual, this can cut the individual off from being recognised by this same agent as having certain other identities (Sen, 2007, p. 62). Moreover, a referent might be put in the right social category by an agent, may be wrongly attributed other characteristics that are normally associated with that category. For example, a referent may correctly be recognised as vegetarian, but based on the same identity, wrongly attributed a love for animals.

Due to the importance of information for the interpretation and understanding of each other, people actively use the sharing of information to signal their identity and engage in distinct social relations and practices (see e.g., Goffman, 1959; Schoeman, 1984). They offer and often even emphasize certain information about themselves that helps establish a certain relation and/or practice, while they repress the information that is irrelevant or maybe even confusing or harmful for this relation. By doing so, it becomes possible for people to establish different kinds of relations and perform different 'roles'. The information sharing practices to distinguish between roles and relations was described by sociologist Goffman in his book *The Presentation of Self in Everyday Life* (Goffman, 1959). Goffman explains these practices by using the theatre as a metaphor: an actor plays a certain role on a stage, and to do so she gives off certain signals to her audience to inform them about the role that she is playing. These signals can consist of information like verbal communication, appearances, body language, attributes and even the selection of a particular environment for the interaction (Goffman, 1959, p. 14). The manner in which personal information is brought to people their attention — the setting, the context, the timing, the method, and the 'who' that reveals the information — therefore all affects how they understand and interpret an individual. Susser thus states: "Drawing epistemic boundaries—determining what people do and don't know about us—is not, therefore, a function of simply concealing and revealing information, but also a function of working to influence how that information is interpreted and understood" (Susser, 2016, p. 3). By giving context to the information, an individual can influence its interpretation. The 'packaging' of information is therefore important to invoke a desired effect on an audience (Susser, 2016, p. 11). In order to make sure that the individual establishes the desired relations and perform specific roles for particular others, it is important for her to *segregate* her audiences so that a specific audience will not perceive an individual in two inconsistent or conflicting roles (Goffman, 1959, p. 137).

In order to make choices on what information to share and to whom, individuals rely on the expectations that they have of the context that they are in. Nissenbaum defines 'contexts' as "structured social settings characterized by canonical

activities, roles, relationships, power structures, norms (or rules), and internal values (goals, ends purposes)" (Nissenbaum, 2010, p. 132). The context in which information is shared determines the significance and meaning of the information (Nissenbaum, 2010, p. 80). The context can also set social boundaries for the kind of information that one is expected to share; not all information is considered appropriate to share in all contexts (Schoeman, 1984, p. 408). For example, people generally share different information about themselves with their partners than with their colleagues. The importance of the 'correct' context is captured by Nissenbaum in the concept of 'contextual integrity' (Nissenbaum, 2004). The contextual integrity of information concerns the reasonable expectations that people have about the norms that govern the information flow in a certain context (Nissenbaum, 2004, p. 137). These are norms 1) about the appropriateness of the information, i.e. norms that "dictate what information about persons is appropriate, or fitting, to reveal in a particular context" (Nissenbaum, 2004, p. 138), and 2) norms that govern the "flow or distribution of information—movement, or transfer of information from one party to another or others" (Nissenbaum, 2004, p. 138). Additionally, it is important that individuals know their audience in order to provide them with the appropriate information and context (Grimmelmann, 2010a, p. 10).

The control that an individual has over the access and perspective of her audiences to and on her informational persona thus plays an important role in the individual's autonomy to give shape to her identity and life (Kupfer, 1987, p. 82). An inability to differentiate between the access that different others have to aspects of one's informational persona can make it difficult, or even impossible, for an individual to play separate roles and engage in various types of relationships (Roessler, 2005, p. 112). Even more, breaches in the contextual integrity of information can disrupt or shatter the self-presentation of the individual in a certain context (Goffman, 1959, p. 63). Once shattered, it will be difficult for an individual to convince a disillusioned audience of the reality of their persona in a specific role (Goffman, 1959, p. 136-137). Take for instance the case of the 'Drunken Pirate', as was briefly set out in chapter 1. The Drunken Pirate photo may lead agents who view this photo to label S with predicates like 'party girl', 'substance abuser','adolescent', 'drunk' and maybe even 'alcoholic'. In the case of S's professional environment, whatever exact predicates her supervisors inferred from the information, they interpreted it in such a manner that they came to the conclusion that she was unfit for the role of teacher — and put an end to S's teaching career. The problem in this case was that signifying objects from a private context spilled over to S's professional context, which shattered S's presentation of herself as a professional and responsible teacher in the eyes of her professional network.

Also, it is important to note that signifying objects themselves can affect our self-perception. Signifying objects can trigger memories or provide us with a view on ourselves that we would not have had without the object. People shape their self-understanding "through the detour of cultural signs of all sorts, which articulate the self in symbolic mediations" (Ricoeur, 1991, p. 79). Signifying

objects that portray us as something, in some context, or as belonging to a particular group, tell us something about who we are. For example, old photographs may remind us of past hobbies, old friends, painful moments, lost dreams, etc. Confronted with this 'old self', we may see ourselves as having succeeded or failed in life, as persistent because we are still chasing the same dreams, or as a lousy friend because we realise that we have not contacted an old friend for ages.

Lastly, it is important to point out that the extent to which the informational persona reflects referents accurately, is not static. As people learn, grow, and experience, their identities evolve over time (Susser, 2016, p. 11). The result is that personal predicates that were accurate at a certain point in the past can become with the passing of time annoying, embarrassing, disastrous or even traumatic for an individual. Think for instance about an ex-partner that is referred to as partner in a particular signifying object. While this information was certainly true at a specific time, and also likely experienced as positive, with changes in the referents lives it becomes inaccurate and maybe even hurtful for the parties involved. If people go through such changes, they tend to approach the period before and the period after the change as as separate periods in time that reflect a change in the self (Bruner, 1994, p. 42). When people relate to this former self from the perspective of their current self-view, they are even inclined to highlight this change and perceive themselves as improved over time (Wilson & Ross, 2003, p. 138-139). With this, people also distance themselves from unwanted past behaviour (Wilson & Ross, 2003, p. 141). With a changed self, individuals can find that certain information does not accurately reflect them anymore, like youthful transgressions, because they do not do such things anymore; they argue that the behaviour belongs to an "old me" (Wilson & Ross, 2003, p. 141). While for the individuals it is clear that the information does not reflect them accurately anymore, they are fearful that this may not be clear, or even not accepted by others (Wilson & Ross, 2003, p. 146). Over the course of their lives people therefore constantly try to update, modify and correct their informational persona so that it presents them as they see themselves to others (Susser, 2016, p. 11).

2.4 Personal information in a technological world

In this chapter, I discussed the main concepts of the analytical toolkit that I will use for the research in the following chapters: the signifying object, the reference, the presence of information, and the informational persona. Signifying objects can contain references to individuals. These references represent a particular piece of information pointing towards a certain subject. All the references referring to a particular person, irrespective of whether these references are correct or incorrect, forms her informational persona. However, agents will never see the complete informational persona. As their access to the informational persona is mediated by signifying objects containing the references, the agents depend on the objects they can access. By being present (or not) for the perceiving agent in

a certain manner, these signifying objects construct a particular presence of the references that they contain. By imbuing certain references with a stronger or weaker presence (or no presence at all), they provide people with a certain view on the subject's informational persona.

This view on the informational persona plays an important role in the manner in which people perceive and interpret each other. While people are likely to have only a partial access to an individual's informational persona, framed by their own perspective, it is in the interest of the individual to make sure that they have access to the appropriate part of the persona from the right perspective. In this, the signifying objects are the revealing mediators, particularly in an online environment. In order to manage their identity for their audiences, individuals therefore constantly need to juggle with the signifying objects that are available in the world in order to reveal certain references, while concealing others. The manner in which the information is made present, the timing, and the who that reveals the information all affect the interpretation of the information by the perceiving agent.

In many cases, these signifying objects are created and brought to our experience by technology. Think for example about photographs, films, online blogs, etc. The technological developments over the last decades have heavily affected the character of the signifying objects and their corresponding presence of information. In order to examine how the Web affects the presence of our informational persona, it is therefore important to first turn our gaze towards technology and get a better understanding of the role that technology plays in the relation between human beings and information.

Chapter 3

Framework part II: technological mediation and personal information

Contents

3.1 Introduction

Signs matter. In the previous chapter I discussed how signs play a fundamental role in our understanding of the world around us, others, and ourselves. However, how these signs appear to us, in what context, and when, is highly dependent on their *matter*, i.e. the material object that reveals them to our perception. Both the information and its carrier enter our perception as one set of stimuli in the form of a signifying object. This makes information in practice often an entity that is affected by its carriers as well as the features of its content. The material existence of the signifying object itself can therefore have far-reaching implications for the perception and interpretation of personal information.

Many signifying objects are created by or dependent on technology. Think about photographs, videos, emails, sms-messages, etc. If we consider writing as a form of technology, it is even difficult to think of a signifying object that is not brought about by technology. The introduction of a new technology can construct a new range of signifying objects and give rise to new aspects of the informational persona. Take for example the photo-camera. The introduction of this technological device had a huge impact on the potential visibility of an individual's appearance (this was especially the case for the portable camera, see Warren & Brandeis, 1890); information that was first solely bound to an individual's physical embodiment, or could only be represented manually by brush strokes or pencil sketches in a time-consuming act, now became possible to capture in a real life representation in minutes, and later in mere microseconds. The photo captures the visible appearance of the individual in a signifying object that is static over time. With this signifying object, the reference to the individual's appearance gains an autonomous presence in the world. This reference allows those unfamiliar with the individual to become familiar with her appearance as it was at the moment the photo was taken.

Technology can thus heavily affect the constitution and appearance of the informational persona. Given this particular role of technology, it is important to get a better understanding of what technology means for the relation between human beings and their world. This chapter delves into this relation. I will show that understanding the manner in which technology can affect our interaction with and experience of information, requires a holistic approach in which agents, technologies, and information are being viewed in the context of each other and the processes that evolve around them.

This relation between humans, technology and information is embedded in the environment; the socio-political system that is our society and shapes the three elements themselves as well as the relations between them. None of these factors can be seen completely separated from the others as they are always affected by them (cf. Stiegler, 2010b). With this deeply entangled triad (or rather tetrad, as all are embedded in the environment) in the back of our minds, I explore in this chapter how technology affects our relation to information and our understanding of the world, and in particular of people. I will do this by first discussing the mediating role of technology and explain how technology expresses a certain

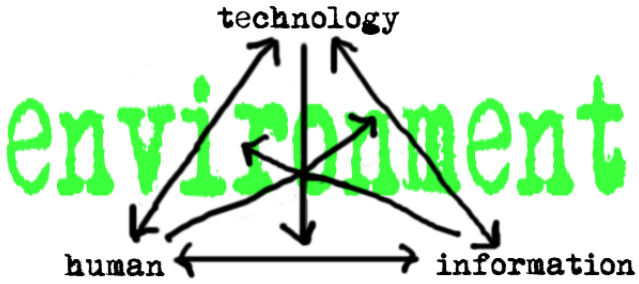


Figure 3.1: Technology as part of a web of relations

intentionality in this mediation. Following this, I discuss the manner in which technology retains information and functions as our 'tertiary memory'. Next, I take a closer look at the necessary process steps for the recollection of information in the tertiary memory; these are the process elements of encoding, storage, and retrieval. After this, we should have an adequate framework to continue our research into the manner in which different online applications affect the presence of personal information.

3.2 Technological mediation

At the heart of this study lies *technology*. Because the particular focus of my research is on digital technologies, with regard to which there is no question of whether these fall under the general concept of 'technology', I will not discuss the scope of the concept of 'technology'. Instead, I focus on the role that technology plays in the relation between human beings and their world. In this section, I explain how technology 'mediates' between human beings and their world and how it expresses a certain 'intentionality' in this mediation. Next, I zoom-in on the user of the technology.

3.2.1 Human-technology-world relation

Technology has been an important part of human life for a long time; society and life as we know would not exist today without technology. Stiegler even argues that technology is a defining characteristic of the human existence because it is constitutive for humanity (Stiegler, 1998). A similar thought we can find with Latour who claims that humans have always been hybrids with technology (Latour, 1993), and Haraway who depict humans as so completely integrated with technology that we have become cyborgs (Haraway, 1991).

When using technology, human beings generally use tools as instruments to reach certain goals. However, the influence of technology goes beyond its instrumental use. Technology can give rise to new experiences by revealing reality

in new ways or provide us with new contexts for our experience (cf. Verbeek, 2005). As such, technology does not only help us to achieve certain goals, it can help us form goals, including new ones, and perform actions that we not only may not have considered possible without the technology, but may have never considered at all. By enabling us to relate to the world in a new manner that is not possible without technology, technology affects our interpretation of the world (Kiran & Verbeek, 2010, p. 418). When we would approach technology as a mere transparent instrument to reach our goals, we thus fail to acknowledge the fact that it actually co-shapes these goals and gives rise to new goals and experiences of the world. Many philosophers in the last century have therefore pointed out that technology is inherently *not neutral*: by creating new options, technology reveals certain aspects of the world while concealing others, it influences human beings, helps to shape and create social identities, enforces power relations, affects culture, production and consumption, and gives rise to occasions of inclusion and exclusion (see e.g., Heidegger, 1977; Marcuse, 1966; Ihde, 1983; Winner, 1989; Latour, 1993; Stiegler, 1998; Feenberg, 2002; Verbeek, 2005; Agamben, 2009).

With the use of technology, we are likely to perceive the world differently (Kiran & Verbeek, 2010, p. 418). We can for instance see bacteria with the help of a microscope, discuss business with people on the other side of the world over the phone, or see an unborn child with an echo device. As such, technology opens up new ways of being-in-the-world for humans by creating new options, giving rise to new practices and by affecting social conventions (Kiran & Verbeek, 2010, p. 415). By allowing us to access the world in a specific manner, technology deeply affects our agency in and interpretation of this world. Technology therefore 'does' something; it intervenes in our world by helping to shape our practices and possibilities (Verbeek, 2005, p. 66-67). It affects and co-shapes both the micro perception, the perception that a single individual has of his or her world, as well as the macro perception, the cultural framework in which technology is used and gains meaning (Verbeek, 2005, p. 172). A valuable concept to introduce here, is the concept of *mediation*. This concept originates from Ihde (see e.g. Ihde, 1983) and was further developed by Verbeek (see e.g. Verbeek, 2005, 2011). It forms the heart of the postphenomenological approach.

'Mediation' entails a description of the manner in which we relate to technology; technology *mediates* our relation to our world. In this mediation, technology takes the form of an active co-shaping of our view on this world; in a certain sense, we see the world *through* the technology. However, the active co-shaping of technology should not be understood in a deterministic manner: technology may trigger certain behaviour, but it does not necessarily cause this behaviour (Hildebrandt, 2015, p. 47). Like information as discussed in section 2.2.2, technology thus 'affords' us things. By being in the world, technology alters the affordances of the world for human agents. As such, technology has a normative impact on the relation between human beings and their world (Hildebrandt, 2015, p. 163). How an agent perceives the affordances of a certain technology, will shape the base for her use of that technology. The perception of the affordances of a certain technology depends on the background of the users, the technology itself, as well

as on the social construction of the technology and its presentation by means of marketing and the like (Nagy & Neff, 2015, p. 6).

By mediating the manner in which humans engage with the world, technologies affect both the manner in which the world and the acting agent is present (Verbeek, 2005, p. 171). As such, the technological mediation is not just an effect between an agent and an object (the world), but consists in a mutual constitution of an agent as a user and the object to which she relates (Verbeek, 2005, p. 130). The user nor her world would be the same without technological mediation, they are in fact the product of exactly this mediated relationship. And because the technologies help to shape our relations to the world including our relation the technologies themselves, technology is only accessible to us in a mediated way. There is thus an intricate relation between human beings, technologies and the world they live in, which all mutually affect each other.

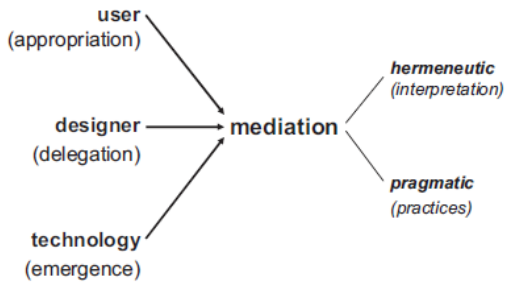


Figure 3.2: Sources and outflows of mediation, (Verbeek, 2011, p. 99)

A helpful diagram to understand the various elements that play a role in manner in which the mediation comes into being, is Verbeek's 'agency and sources of mediation', reproduced in figure 3.2. Here we can see that the user, the designer as well as the technology itself help to constitute the mediation (I will get back to the relation between the technology and the designer in the next subsection). While the user plays a role in the establishment of the mediation, the technological mediation in turn affects how the user experiences the world and engages with it by affecting her perception, interpretation, and practices. Technology thus mediates human experience on a hermeneutic as well as on a pragmatic level (Verbeek, 2011, p. 99).

3.2.2 Intentionality

When a human agent focuses on the world while using a technology, her intentionality is mediated by this technology (Verbeek, 2005, p. 116). However, as I will explain in this subsection, in this mediating role, the technology itself also expresses a certain 'intentionality'.

In their mediating role, technologies help structure and organise our world

(Kiran & Verbeek, 2010, p. 417). By using technologies we interrelate to our environment in a specific manner; the technology shows certain aspects of a technological co-shaped reality and makes certain things stand out, while often at the same time obscuring or reducing the presence of other elements (Verbeek, 2005, p. 131). Think for instance about the use of a telephone: it makes sound — a voice — stand out, while obscuring the rest of the individual at the other side of the conversation. The technology thereby has a certain directionality towards reality (Verbeek, 2005, p. 114). Additionally, technologies have a certain directionality "within which use-patterns take dominant shape" (Ihde, 1990, p. 141). Verbeek phrases this as a "a trajectory that promotes a specific kind of use" (Verbeek, 2005, p. 115). Due to this directionality, technologies "*suggest, enable, solicit, prompt, encourage, and prohibit certain actions, thoughts, and affects or promote others*" [emphasis original](Lazzarato, 2014, p. 30). The directionality of a technology thus affects the intentions and views of its users. It establishes a particular relation between the user and her world (Verbeek, 2005, p. 115). The directionality is shaped by the concrete design of the technology, its materiality.

By giving a technology certain material properties, the designer of the technology aims to give it a certain directionality. Additionally, the designer (or seller) will generally give clues to users on how to understand and use the technology by means of instruction manuals, using signs in the design, and by marketing the technology as a particular technology. However, while the designer determines the material properties and promotes a particular use context, this does not necessarily lead to a use and consequences of the technology that the designer intended. Often, technologies have unforeseen side-effects, or people can willingly look for ways to use technologies in a manner different from the use that was intended by the designer.[1] As such, technologies have a certain autonomy in which they are present for agents; the technology presents itself and not necessarily conform the plans of its designers (Chabot, 2013, p. 15). In its autonomous existence, the technology expresses a particular directionality towards reality as well as towards its way of use. I understand this directionality as the materialised inclination of the technology towards highlighting certain appearances and realising certain affordances above others for human beings. This directionality of technology is what Verbeek calls its 'intentionality' (Verbeek, 2005, p. 115). With this concept Verbeek aims to capture two meanings of intentionality in relation to technology: "a first referring (...) to the 'intentions' of the technology itself, the second (in the more general phenomenological sense of 'technologically mediated intentionality') to the relations between human beings and world that are mediated by the technology" (Verbeek, 2005, p. 116). Technological intentionality does not mean that a technology can form an intention like a human being can form an intention, for that would require technology to be conscious. Instead, 'intentionality' in this sense should be understood as a directedness towards something: "The intentionality of artifacts is to be found in their directing role in the actions and experiences of human beings. Technological mediation therefore can be

[1] See e.g., groente, "Philosophy of hacking", *PUSCII blog*, 2014. `http://www.puscii.nl/blog/content/philosophy-hacking`, last accessed 30-10-2019.

seen as a distinctive, material form of intentionality" (Verbeek, 2011, p. 57). However, as technologies always play a mediating role and are dependent on the human intentionality supporting their use, the intentionality of the technology is necessarily part of a hybrid affair of the technology and its users (Verbeek, 2011, p. 58). As part of this hybrid affair, the intentionality of the technology does not determine how a human agent uses or perceives the world through the technology, but it does help to co-shape the intention of the user (Verbeek, 2011, p. 58).

3.2.3 Users

As became clear in the previous sections, the technology is inherently intertwined with its users. These users, however, generally do not comprise of the whole human species, but are restricted to a certain user group, which is co-shaped in turn by the technology. A technology sets certain boundaries to its use. By requiring certain skills and resources, it can include and exclude people from its use, and can give rise to different sorts of relations to the technology (Oudshoorn & Pinch, 2003). These different relations go further than relations of use, non-use, and divergent use of the technology; technology can also give rise to different social relations between users and their world. For example, the use of a particular technology can imbue users with a certain social status, (Oudshoorn *et al.*, 2004, p. 40). As such, technologies can "act as sources and markers of social relations and can shape and create social identities" (Oudshoorn & Pinch, 2003, p. 12). Technology can even actively socially frame its users by addressing them as a certain category or group and establish certain norms and expectations in their users (Stanfill, 2015, p. 1064). A famous example of this is the inscription of gender in electric shavers (see e.g., Oudshoorn *et al.*, 2002; van Oost *et al.*, 2003). By bringing on the market an artifact with the same function, but in two different design-styles, Philips suggests in their technology design the existence two different types of users, namely on the one hand a user that prefers dark coloured and right-angled items that can be used on many locations and allows the user to tinker with the artifact, and on the other hand the 'lady' user, a user that prefers pastel colours, rounded angles, and has no interest in accessing the technology of the artifact (Oudshoorn *et al.*, 2002, p. 475). While users are free to choose which shaver they use (or none at all), "the gender script of the Ladyshave inhibits (symbolic as well as material) the ability of women to see themselves as interested in technology and as technologically competent, whereas the gender script of the Philishaves invites men to see themselves that way. In other words: Philips not only produces shavers but also gender" (van Oost *et al.*, 2003, p. 207).

3.3 Information and technology

In the previous section I discussed how technology mediates our relation to the world, and how it expresses a certain intentionality therein. In this section, I add information to the mix. I will first delve into the manner in which technology

mediates our relation to the world by materialising information. Following this, I discuss the three process steps that we minimally need in order to be able to interact with personal information mediated by technology.

3.3.1 Retention of information in technology

In order to take a closer look at the relation between technology, human beings, and information, I start by discussing Stiegler's work on the functioning of technology as a 'tertiary memory'. His theory is valuable to include here because it provides insight into this relation. Stiegler explores the manner in which technology retains information for human beings. When information is materialised and spatialised by means artefacts and techniques, it becomes an exteriorised memory that is easily "transmissible, inheritable and adoptable" (Stiegler, 2011, p. 117) and "cumulative" (Stiegler, 2009, p. 4). Stiegler captures this technologically retained information in the notion of 'tertiary memory', which he bases on the work of Husserl.

The tertiary memory is the third of three types of information retention; the others are the primary and secondary memory (Stiegler, 2011, p. 111-112). As I will explain, these three tie closely to the human-information model presented in chapter 2. Adding Stiegler's theory to this model allows us to gain a better insight in what the material character of signifying objects means for the presence of information and the interpretation of the world by an individual.

The **primary memory**, also called 'primary retention', is the individual's experience of the present (Husserl, 1991, p. 32).[2] This memory is by the individual perceived as a continuous singular experience (Stiegler, 2011, p. 111). The primary memory roughly corresponds with the deriving of information from stimuli by the individual (which already requires a certain interpretation) in the human-information model discussed in section 2.2.2.

The primary memory does not yet involve a recollection of this experience. However, it does already entail a *selection* of what is retained in the continuous experience of the present (Stiegler, 2014, p. 52). This selection is based on an individual's knowledge background and experiences which the individual gained prior though the primary memory (Husserl, 1991, p. 37). The retention of these earlier experiences is the individual's **secondary memory** (Stiegler, 2011, p. 112). The secondary memory functions as a conceptual filter discussed in section2.2.2 as it shapes the selection criteria for the primary memory.

Both the primary and secondary retention of information take place within a single individual. However, human beings also retain information outside themselves. This external information retention is the **tertiary memory**. The tertiary retentions are materialised and spatialised secondary retentions that are encoded into artefacts and techniques (Stiegler, 2011, p. 112). The information has been given shape in the outside world by being materialised in an object. In this form, it can be 'recollected' and interpreted because it gives off certain

[2]Husserl's concept of 'primary memory' (also called 'fresh memory') is synonym to the term 'retention' which he uses later on in his work (Husserl, 1991, see p. XVI and 32).

stimuli. The materialised information is in this sense a memory *exterior* to the individual: it is not biologically given, but supported by exterior objects and factors (Stiegler, 1998, p. 57). For example, a secondary memory is exteriorised by writing down an experience in a diary.[3] With this, the secondary memory becomes a tertiary memory in the diary. However, also the diary itself embodies particular references. For example, it can signify its meaning and use to a particular beholder. Tools themselves thus also contain certain references (Stiegler, 1998, p. 254). Stiegler therefore argues that the tertiary memory is not restricted to objects with their roots in information technology: instead, he argues that all technology retains information outside of human beings.[4] This is the point where I suggest to make a little sidestep and bridge between Stiegler's concept of the tertiary memory, and Verbeek's notion of technological intentionality. Stiegler argues that all technology retains information, because it has a certain material existence for human agents. In a similar line, we saw how Verbeek recognized a certain technological intentionality exactly in this materialised design of technology. If we combine these two perspectives, we can see that all technologies express a certain intentionality in the information that they convey due to their material properties and characteristics. This means that when an information technology conveys some information that we send with it, it never just presents the information we send, but also always at the same time the information that is an inherent affordance of the technology itself. As the total is presented to an agent's perception as one set of stimuli, the information that she receives is necessarily co-constituted by the mediating technology. For instance, if I send my mother the message "Thanks for the flowers!" through the Signal application, she perceives not just the message, but the text as displayed in the interface of the application as well as of the device, which next to the text also reveals my name, a profile photo, the previous messages, maybe the names of others, the date and time, the design of the device itself, its social use, etc.

Generally, the tertiary memory has a collective nature, because it is accessible to multiple people. By using the same tertiary memory, individuals share their knowledge and experiences with others — ranging from close others to those unknown and/or at distance in space, or in the future. Such a collective memory influences what a group of people remembers and what they believe to be true (Wegner, 1987; Sparrow *et al.*, 2011, see the initial reseach of Wegner on the 'transactive memory', and his later view that this is also applicable to technology). As the tertiary memory tends to extend the presence of information in space and time, it affects agents on a perceptual and/or actional level (Ross, 2013). For example, materialised past experiences allow agents to experience "a morsel of time' (...) in the present" (Middleton & Brown, 2005, p. 149). Here, it is important to note that individuals will very likely not have the (exact) same recollection when they encounter a particular part of the tertiary memory. For instance, family

[3]This exteriorisation goes hand in hand with the interiorisation of the technical skills by the individual (Lemmens, 2015, p. 348).

[4]A similar thought can be found in Flusser's work who describes how information is stamped into leather when it is used to produce the "cultural object 'shoe'" (Flusser, 2011, p. 108).

members may recall different experiences when viewing the same holiday photo (Middleton & Brown, 2005, p. 144). This has to do with the distinct personal (interpretive) background that every individual necessarily has (see section 2.2.2).

Lastly, the tertiary memory is not just a recollection. It forms an inherited past into which humans are born (Stiegler, 2011, p. 112). As we are born into a world with technology all around us, we are born into these memories, practices and systems. It is part of our cultural memory and forms a base for our expectations of the future. The tertiary memory therefore constitutes a *protention*: an anticipation for the future (Stiegler, 1998). Despite the fact that individuals have not experienced the content of the tertiary memory for themselves, it thus nevertheless shapes the normative and experiential backdrop of their experiences. As such, it influences people their experiences, choices, behaviour, and expectations. The information that is present in the tertiary memory therefore has a significant impact on the manner in which human beings understand and experience their world and their culture, as well as their own agency and identity (Brockmeier, 2002, p. 26).

3.3.2 The tertiary memory as a process

The technological constitution of information not only impacts *what* that is retained in the tertiary memory, but also *how* we use it. As we interact with external information collection and processing, we often experience new and different information flows compared to what the 'naked' human body would be capable of. In his book *Natural-Born Cyborgs*, Clark therefore states that humans are "products of a complex and heterogeneous developmental matrix in which culture, technology, and biology are pretty well inextricably intermingled. (...) Ours are (by nature) unusually plastic and opportunistic brains whose biological proper functioning has always involved the recruitment and exploitation of nonbiological props and scaffolds" (Clark, 2003, p. 86). As such, technology and human beings influence each other and co-constitute each other by being part of one hybrid functional system (Heersmink, 2012, p. 122-123).

On the information level, technology allows us to deal with complex problems and a magnitude of information by giving us tools to store, alter, combine, and transform information in ways that would require a lot of time and energy from our biological brains — if they even would be able to process it at all (Clark, 2003, p. 78). Especially learning to read and write allowed human beings to overcome storage limitations of the biological brain and thereby highly affected our knowledge and consciousness (Wolf & Stoodley, 2008, p. 216-217). By externalising information, we are able to transfer the processing of information and the 'burden of 'remembering' to artefacts (Middleton & Brown, 2005, p. 162). This use of technology has many advantages, but it comes at a cost: we become dependent on the technology. Plato already expressed criticism towards the cost that comes with the unburdening of the brain by means of reading and writing. Via the character of king Thamus he states that those who acquire reading and writing "will cease to exercise their memory and become forgetful; they will rely on

writing to bring things to their remembrance by external signs instead of on their own internal resources" (Plato, 1973, p. 96). In their unburdening the biological brain, technologies compensate for human flaws by protecting against forgetting, but at the same time they deepen these flaws by alleviating the need to remember everything, and consequently, they diminish the need to train and improve our memory. Once we know that certain information is easily accessible, we take it for granted that we can retrieve the information from our technological environment, and we tend to use the capacity of our brains to help us remember how to find things, instead of recalling the things themselves (Sparrow *et al.*, 2011). In this light Stiegler, in the footsteps of Derrida, argues that technology is a 'pharmakon', a poison that is at the same time its own antidote (Derrida, 1981; Stiegler, 2012). Take for instance the use of an agenda. By writing appointments that we want to remember down in our agenda, we relieve our brain of the burden of having to remember. Hereby we 'poison' our brain by allowing it to forget instead of training it. However, the agenda itself functions as a remedy to this forgetting, because if we consult it, we will recall what our brain did not have to remember. With this, we come to rely on technology: our remedy which is at the same time our poison.

Given this impact of the tertiary memory on how we interact with information, it is important to take a closer look at the manner in which the tertiary memory functions. For this, we can find some helpful anchor points if we approach the tertiary memory as a *memory process*. The tertiary memory process consists of at least three elements that are fundamental to every memory system:

> *Any memory system — whether physical, electronic, or human — requires three things, the capacity to* encode*, or enter information into the system, the capacity to* store *it, and — subsequently — the capacity to* retrieve *it* [emphasis original] (Baddeley *et al.*, 2009, p. 5).

These three elements interact and shape the memory process: the manner of encoding determines what and how something is stored, which in turn determines what can be retrieved (Baddeley *et al.*, 2009, p. 5). The concrete functioning of these three elements can differ per type of memory, but they all necessarily comprise the same three elements. In the tertiary memory the processes of encoding, storage and retrieval have a form that is external to the human agent. In order to get a better grip on these process elements, I will discuss here what they mean for the tertiary memory.

3.3.2.1 Encoding

In order for information to be retained in the tertiary memory, it first needs to be externalised: the information needs to be fixated in an exterior carrier. This is done by *encoding* the information into these carriers. With the encoding of personal references into the tertiary memory, the content of the informational persona is created.

When references are encoded into objects, they are given shape in a certain

format. Following the Article 29 Data Protection Working Party (hereafter: WP 29), I take the format to both include the form of the information, which can be for example "alphabetical, numerical, graphical, photographical or acoustic"[5], as well as the carrier of the information, like the paper, cassette or computer that retains the image, text, sound, etc.[6] The format needs to be based on a common code that allows at least a partial convertibility of the human memory into the tertiary memory (Hui, 2016, p. 319) — it is thus an *encoding*. As such, technologically created images are not mirror images, but representations (in Flusser's terms: projections) of something (Flusser, 2011, p. 66). Different formats can project the same referent. See for example figure 3.3, which shows some examples of references in various visual formats to the referent 'cat' as the animal in general. The format of the reference can divert from the informational content. For instance, we can use textual language to describe a visual appearance: "the cat is black, has green eyes and long whiskers".

Figure 3.3: The reference 'cat' encoded in different types of formats

The process of encoding plays a pivotal role in the specifics of the particular representation that is created of an event or person *as a reference*. First of all, what is encoded is never "the event as event" (Ricoeur, 1976, p. 27). Instead, it is always a selection converted into a certain format. This selection is never a reference to what actually happened, but a selected framing thereof (Stiegler, 2009, p. 115). The selection of what is encoded is in itself therefore also a forgetting; what is selected is remembered, and what is not selected is lost (Brockmeier, 2002, p. 22). In this process, the referent is thus reduced to a particular reference. Moreover, by selecting this particular reference to encode, and not others, the reference is given a certain importance as it is considered meaningful enough to retain (Stiegler, 2009, p. 115). By being encoded, the reference thus gains a certain presence on the quantitative as well as the qualitative level.

In the encoding process, technology plays an important role; the technologies we have at hand, and the effort and skills they require, affect the selection of what is encoded as well as the format of the encoded reference. Ihde therefore describes encoding by means of writing as "technologically mediated language" (Ihde, 1990, p.81). Let me explain the impact of the technological mediation on the encoding

[5]WP 29, Opinion 4/2007 on the concept of personal data, p. 7.
[6]Ibid., p. 7.

by means of an example. Imagine that we see a poster of an art exhibition that we would like to visit. Because we are overwhelmed with work, we do not trust our own brain to remember the address and duration of the exhibition. We therefore seek the help of tools to retain the information. If it turns out that we only have a photo camera at hand, we are restricted to visual encoding by means of images and would therefore likely make a photo of the poster. This results in a colourful visual representation that allows us to see the poster, the text on it, and maybe even part of the environment where we encountered the poster. This would be different if we only had access to a pen and a notebook. Unless we have spectacular drawing skills, many of us would prefer to materialise the information by means of plain text. This would lead to a short textual representation that only reveals the name of the exhibition, its address and duration. Thus while we express a certain intention in the encoding by means of selecting and encoding a signifying object by copying a part of the information presented by the poster, the technology we use also expresses a certain intentionality by only affording particular ways of encoding and requiring certain skills, while it hampers, or even prohibits other ways of encoding. The encoding process is therefore a hybrid action in which human and technological intentionality are intertwined in the materialisation of a reference.

Lastly, it is important to note that the encoding process affects the relation between the encoder and the encoded. By externalising thoughts and experiences, a certain distance is created between the encoder of the thought and the materialised information (Ricoeur, 1976, p. 36). As the object and the author are separately existing entities, "the author's intention and the meaning of the text cease to coincide" (Ricoeur, 1976, p. 29). The human as the agent that conveys the message disappears. Instead, "material 'marks' convey the message" (Ricoeur, 1976, p. 26). The signifying object therefore receives a certain 'semantic autonomy' (Ricoeur, 1976, p. 29). This results in a distantiation between the author and the content about which Hildebrandt states: "This distantiation is afforded if not imposed by the material inscription, fixation, externalization and objectification of human thought, which is — in turn — co-constituted by this externalization and distantiation" (Hildebrandt, 2015, p. 48). The prospect of encoding influences the manner in which we think about what to encode; when we know we are going to put a particular thought on paper and will be at a certain distance to it, we already anticipate on this external perspective when we form the thought (Hildebrandt, 2015, p. 48). As such, the experience of the encoding process affects the content that we are likely to encode.

3.3.2.2 Storage

By means of encoding, information is materialised in a certain object. This materialisation gives shape to the presence of the informational persona. By being encoded into an object, the reference is freed from the limits of 'situational reference'; the reference can exist separate from that to which it refers (Ricoeur, 1976, p. 36). However, this existence is now tied to the object into which it is

encoded.

By being encoded into a particular object, the information takes on the properties afforded by the object's storage capacities. The object generally allows a relatively stable retention of information as "the object lends something of its material durability to [that which] we wish to recall — it projects something of its stability into the fluidity of our past experience" (Middleton & Brown, 2005, p. 150).[7] When storing information in a particular object, the information inherits several properties from the carrying object that are worthwhile to take into account when we assess the impact of the tertiary memory on the storage of information. When researching literature and technology, I identified five main main properties that play a significant role: (1) the types of information that can be stored, (2) the quality of the stored information, (3) its flexibility, (4) the quantity that can be stored, and (5) the time that the object can be retained. I will discuss these properties subsequently.

Let us start with the types of information that can be stored. The information needs to be encoded and stored in a certain object. However, the objects in which we store information come with certain restrictions with regard to the type of information that they can hold and what kind of encoding techniques need to be used for this. Depending on the object, information can be stored in a visual, audio, or touch-related form. For example, a piece of paper cannot store a song as sound, but it can store the song in a printed or Braille form of musical notation. By being able to carry only certain types of information, the objects themselves thus leave a strong mark on the types of information that we can find in the tertiary memory.

The second property that I will discuss is the quality of stored information. This quality can be viewed from two (often intertwined) perspectives, depending on one's understanding of the term 'quality'. On the one hand, the 'quality' can refer to the material characteristics like the level of detail and lack of decay of the stored information over time. The better the signifying object is preserved in a state resembling the state when the information first was encoded, the higher its quality. On the other hand, 'quality' also can be taken as a marker for the accuracy of the information represented by a signifying object. Information can be considered to be of a good quality if the information is truthful, accurate, relevant, detailed, meaningful, scientifically proven etc. This is where we can already see a potential friction between the references that the information contains, and the representation of the information by a signifying object; the durability of the material can negate the accurate quality of information as the passing of time is often a factor that diminishes this accuracy. An example that shows this, is a passport photo. These photos are accurate at the time that they are taken and are therefore considered appropriate for identification purposes. However, as we physically get older and start to look different from our younger selves, old passport photos tend to lose their accurate representative value. Meanwhile, these photographs themselves can be in a pristine condition and still be of high quality

[7] Though the duration of this storage can be short. Flusser for instance, already considers the oral transmission of information as memory locked in airwaves (Flusser, 1990, p. 397).

on the material level after fifty years.

Thirdly, the flexibility of the signifying object is important to take into account. The flexibility of stored information is the degree to which the information can be adjusted over time. Information stored in physical books is for instance relatively inflexible; once printed, the text in a book cannot be altered, except by manipulating the carrier by means of addition (e.g., by writing with a pen in the book) or by destruction (e.g., scratching in the text, tearing out a page). Contrarily, digitally stored information is relatively flexible. I will discuss this in section 4.2 of the next chapter.

Fourthly, the quantity of information that a particular technology can store matters. The quantity of the storage concerns the amount of information that can be stored in a specific region of the tertiary memory. This highly depends on the characteristics of the objects that are stored. For instance, due to the physical properties of books, there is a maximum number of books that people can store in their houses. At a certain stage, they will run out of storage space. At this point the agent will need to either not acquire new signifying objects, to discard (some of) the objects that she currently has, or to buy a bigger house or rent a storage locker. Such volume limitations necessarily lead to a form of 'forgetting-by-selection' with regard to the tertiary memory in question; either the new or the earlier retained is removed from the tertiary memory. In many cases, this will mean that outdated signifying objects, or those with little meaning to the agent controlling them, will be discarded to make room for more meaningful and/or contemporary information.

However, human evaluation is not the only force that plays a role in the retention of information over time. This brings me to the fifth property: the time that the signifying object itself can be retained. The material character of the object often imbues the information with a certain lifespan, but also a certain fragility. The durability differs per object. Some signifying objects have a very short lifespan, like a message written on the beach during low tide, or a self-deleting digital message. While most objects have a relatively durable character, their lifespan can be shortened by various factors. Fires, leaks, and natural disasters can reduce or even end the time that a signifying object can be retained. Also, as technological developments are ongoing, especially on the level of digital technology, the hardware (e.g., the shift from floppy disks to diskettes to USB sticks) and software (e.g, the shift from Word Perfect to Word) go through cycles of innovation. While this strictly speaking does not affect the retention of the stored content, it does affect the access to the content. I will discuss this in the next subsection.

The particular storage properties of signifying objects heavily affect the presence of information over time. This, in turn, affects the manner in which individuals experience 'their' history; due to their durability, signifying objects may be granted a certain authority as a 'true' representation of the past over time (Middleton & Brown, 2005, p. 175). Especially since human memories are not collective and generally fade, the information that is retained in the tertiary memory is likely to become authoritative for the understanding of our world.

3.3.2.3 Retrieval

The materialisation of a reference affects its character and relation to time and audiences (Ricoeur, 1976, p. 35). Once information is stored unto a material carrier, individuals distanced in time and space from the original encoder can take notice of the information. However, in order for individuals to access the information, the information needs to be retrieved. This retrieval process is therefore a vital part of the tertiary memory.

The manner in which information can be retrieved affects the audiences of the information, as well as the setting in which the information is presented to individuals. It plays an important role in the presentation of the informational persona and the composition of its audiences. In turn, how the information can be retrieved and used, and by whom, depends on technological character of the signifying objects into which information is encoded. The object shapes the information's compatibility for certain information technologies, ways of transmission, and manner in which it can be perceived. Each technical object "has its own material limits and resistances, and these dictate what humans can achieve when they are connected to such artefacts. (...) The material limits will be different for each technology" (Barnet, 2013, p. 52). The object sets the retrieval requirements for users, like the needed devices, resources, and skills. For instance, information stored in purely physical carriers, even if theoretically publicly accessible, sets physical limitations to the access of this information (Nissenbaum, 2010, p. 54-55). Additionally, the individual herself plays an important role in the retrieval process, because she is the one who chooses the method to retrieve information. For example, an individual that wants to retrieve certain content from a library may choose to look at the books to find what she wants, or she can choose to make use of the library's index system. These methods will likely lead to the retrieval of somewhat different content.

With ongoing technological developments, the retrieval of information in the tertiary memory can be increased, as well as reduced. On the one hand, as technology evolves quickly, devices and formats may become outdated and thereby increasingly difficult to access. Take for example the Betamax tape, mini-disc, and the laser disc. The content on these devices needs to be retrieved with devices that are increasingly difficult to come by. The result is that, despite the fact that the signifying objects are successfully stored, they are difficult to retrieve and are thus less likely to reach an audience. As such, references can lose their presence as the signifying objects become part of an outdated technology. On the other hand, the shifts in popular technologies may activate users to engage with the content. As the shifts in popular technology forces people to reformat their encoded memories into the new medium, the users are likely to revisit their old content and select what to reformat and what to discard. This brings the information again to the awareness of the user. There is therefore also some element of increased presence of information due to the changes in storage formats.

Furthermore, technological developments can also greatly enhance the scope of the retrieval. With technology we can transport information across vast distances

in mere seconds, or recover content from a long ago past. With this, the segregating impact of spatial and temporal distances is reduced. This is a phenomenon known as "time-space compression" (Ross, 2013). With this space-time compression, the retrieval of information imbues a reference with a presence in the here and now for a retriever, despite the fact that the signifying object originates from the other side of the world or from a different era. As such, the act of retrieval allows for the spatially and temporally far away to become a part of the present of the receiver. The retrieval of signifying objects can thus 'actualise' the past and the far away in the here and now (Middleton & Brown, 2005, p. 164). Yet, this actualisation of the past and the far away is co-shaped by the retrieval process itself, because the retrieval itself entails a selection (Brockmeier, 2002, p. 22). That which is selected is made present, while that which is not selected remains in oblivion. An example that clearly shows this, is the retrieval of information by means of a search system; based on a keyword, such a system retrieves a particular set of results from a database. The references brought forward by the returned results gain a certain presence, while those that do not make it into the search results remain out of sight. With this, the retrieval process is 'discriminatory'; "any cue to recall, whether self-initiated or externally initiated, defines an item or set of items to be discriminated from possible competitors and retrieved" (Bjork, 1970, p. 255).

Technologies that are specifically used for information retrieval play an increasingly important role in the contemporary use of the tertiary memory. Due to the amount of information available, the human memory cannot keep track anymore of all the externalised information and has to rely on index systems and the like (Leroi-Gourhan, 1993, p. 262-263). By mediating the retrieval, these technologies affect the manner in which humans beings perceive their world. An example par excellence of the increasingly important role of technology in information retrieval, is the use of search engines by Web users; many users depend in their online information retrieval on these technological mediators. I will discuss this in full in chapter 6.

3.4 Going online

In this chapter, I discussed that technology is inherently non-neutral. It gives rise to new ways of perceiving the world around us and offers us new goals that did not exists or were impossible without technology. Connecting to the work of most notably Verbeek, I discussed that technology has a certain directionality in the manner in which it establishes a particular relation between the user and her world, and offers her certain perceptions and goals. This directionality is embodied in the concrete material design of the technology. While the technology is shaped by its designers, its use and effects are not limited to their intentions. Instead, the material form of the technology has a autonomous existence which itself expresses a distinctive directionality that directs the experiences and actions of users towards something. This directionality of technology is a material form

of 'intentionality'. However, as technologies always play a mediating role and are dependent on their human users for the manner in which they are actually used and have effects, the intentionality of the technology is necessarily part of a hybrid affair of the technology and its users. As part of this hybrid affair, the intentionality of the technology does not determine how someone uses the technology, but it does co-shape the user's intention. This hybrid intentionality, and the respective weight of the technology and the human agent in the forming of this intentionality, constitutes one of the crucial concepts in this study.

From there on, I added 'information' to the mix. For this, I connected to the work of Stiegler, and to a lesser degree Clark, and examined how the non-neutral and co-shaping character of technology constitutes an exteriorised memory that mediates our relation to the world and expresses a certain intentionality herein. When we interact with information in this tertiary memory, technology affects this interaction on the three process levels of encoding, storage and retrieval. The process of encoding is hybrid affair that already deeply impacts what is retained in the tertiary memory, as well as our relation to it. By being encoded and stored in an object, the information gains a certain semantic autonomy that does not necessarily coincide with the intentions of its author. On the storage level, we can see that the object imprints the characteristics of its material form on the information that it contains; the information needs to be given a certain material shape and inherits the durability of the object. Lastly, technology highly affects the conditions of the retrieval of information. This affects the presence of a particular piece of information compared to other information, as well as the potential audiences of the information.

Taking all this into account, we can see that technology deeply impacts our relation to information and the manner in which we perceive the world around us. This impact includes the manner in which we perceive others, as well as ourselves. By mediating personal information, technology can present us a certain view of someone and thereby affect the way in which we understand their identity.

One of the technologies that mediates personal information is the Web. The manner in which personal references are made present by the Web is at times experienced as problematic. The sense of a pressing problem was apparently strong enough to drive the EU legislator to develop a right that should address certain instances of the availability personal information in the online realm — art. 17 GDPR. Art. 17 GDPR is presented as a 'solution' to problems that individuals experience as a result of the availability of their personal information on the Web. However, what exactly the problem is, and whether art. 17 GDPR can resolve it, is not yet clear. In order fill this knowledge gap, I will research the problematic impact that the Web has on the presence of personal information in the following chapters. However, because the Web as a singular case study is too broad, I split this research into four sub-case studies of technological mediation; three online applications and one phenomenon. In the following four chapters I will discuss (1) basic web pages, (2) social media, (3) search engines and (4) online virality, while using the two framework chapters as an analytical toolkit.

Chapter 4

Web pages

Contents

4.1 Introduction

When the Net absorbs a medium, it recreates that medium in its own image. It not only dissolves the medium's physical form; it injects the medium's content with hyperlinks, breaks up the content into searchable chunks, and surrounds the content with the content of all the other media it has absorbed. All these changes in the form of the content also change the way we use, experience, and even understand the content.

Nicholas Carr, *The Shallows*, 2010

The World Wide Web. I doubt it needs any introduction in 2019. As an application of the internet, the Web became available around 1995 for Western society at large (Castells, 2002, p. 17). As the Web became increasingly popular with the general public, the technology was quickly adopted in the daily routines of people and became what Silvertone and Haddon call 'domesticated' (Silverstone & Haddon, 1996). Now, billions of people have access to the online world and interact with it daily on smartphones, laptops, tablets and desk top computers. Due to the Web's worldwide implementation and use, it has made a tremendous impact on the availability of information, including personal information — and with that, on our informational personae. This online personal information can represent the referent in unforeseen and unwanted manners. A referent who herself added her personal information to the Web explains:

When I was 20 years old, I made a website for a college course about building a digital identity. Today, it makes me cringe—largely because the site has become such a stubbornly resilient piece of my digital identity. At the time, I was proud. In a matter of weeks, I had learned to cobble together a series of letters and symbols into a code that'd transform into a real, live website for readers everywhere. But seeing it 10 years later is like looking back at embarrassing old family videos, pondering why you would ever say or wear what you did.[1]

The goal of this chapter is to examine how the Web affects the presence of personal information for users, and why this may cause problems for the referent. In order to get to the bottom of the manner in which these problems come into existence, I cannot restrict this research to the experience of the user. I will also need to look at what happens behind the screen in order to examine the roots of the problems, the technological intentionality herein, as well as contemplate the hermeneutic challenges that this mediation brings forth with regard to users. However, the focus will remain on that which is perceivable to common users. I

[1] Kaitlin Mulhere, "An Embarrassing Website I made in College Has Followed Me for a Decade. Here's How I Finally Erased It From My Google Search Results", *Money*, 2018. `http://money.com/money/5441177/manage-google-results-online-reputation/`, last accessed 25-04-2019.

will not discuss the presence of personal information for the controllers of websites, who can access these sites from the 'backstage'. Also, as already pointed out in section 1.2.1 of chapter 1, I will not discuss the actions that users take based on their encounter with the personal information, like firing the referent, denying her a particular service or ending their relationship with her, although I will hint at such consequences now and then by means of examples.

Due to the extensive scope of the Web as a case study, I have chosen to break the analysis of the Web's impact on the presence of personal information down in more digestible parts (see section 1.3.2). This chapter will be the first of four case study chapters, which together form the foundation for the assessment of art. 17 GDPR's functionality in chapter 8. For the current chapter, I have made a somewhat artificial split in the online information sources and will focus solely on the Web in its most simple form, namely basic websites, without looking further into specific and more complex web applications. Particular types of applications and internet-based services with a web interface, like social media sites and search engines, will be addressed in the upcoming chapters. With this artificial split, I aim to trace the different elements that play a role in the problems raised per technological application. Because all the following chapters see to particular websites or online phenomena, this chapter serves as the base analysis on which the following chapters build forth. The focus of this chapter will therefore be on what it means for the informational persona when a signifying object is 'assimilated' by the Web. Because a web page is not a single technology, but a set of highly intertwined and layered technologies which all have their particular affordances, I will start my inquiry at the base, by first exploring the characteristics and affordances of digital information. From there on, I trace the impact of the technological mediation on the online assimilation of personal information in three directions: the production of personal information (and thus the content of the informational persona), the presence of information, and the composition of its publics. These three traces link to three main elements that shape the perception of the informational persona, namely the content of the information, its presence and its audience (see section 2.3). Lastly, I will conclude this chapter by reviewing how the assimilation of personal signifying objects by the Web can complicate the portrayal of an individual by her informational persona.

4.2 Interfaced objects

Digital objects have a peculiar character, which affects what we can do with them, as well as how we experience them. In this section, I discuss the general affordances of digital information. This will be followed by a discussion of the manner in which digital objects become present for our perception.

4.2.1 The affordances of digital objects

Personal signifying objects on the Web are 'digital' objects.[2] Before delving into the impact of the Web on the presentation of our personal information, it is therefore important to first take a closer look at what it means for personal information to be digital. In this section, I will examine the implications of digitising signifying objects, and already briefly touch upon the implications of their online assimilation.

In its core, a digital signifying object is the encoding of an informational unit into a discrete set of binary values. These binary values are expressed in ones and zeros, the 'bits'. A bit is "the smallest amount of information a computer can store. Think of a BIT as a switch that is either 'on' or 'off'. When a BIT is 'on' it has a value of 1; when it is 'off' it has a value of 0" (Commodore Business Machines, 1982, p. 76-77). Various types of information, like sound, text, images and video, can be encoded into such a set of discrete values. Yet, in this process of digitisation, the computer impresses certain characteristics of the digital upon the information that it assimilates. It is important to note that in the case of some signifying objects, this digitisation may be a longer process, influenced by multiple devices. Take for instance a photo. If a photo is taken on a mobile phone, the digitisation is instant and takes place in one device. However, if the photo is taken by an analogue camera, the camera itself impresses certain analogue characteristics on the object, like sharpness and granularity. However, in order to be digitised, the photo will need to be scanned. This, in turn, will impress the characteristics of the device on the object by translating it into a digital image with a certain resolution and color style. The representation of the information in the final digital object that is stored on the computer is thus already heavily influenced by the devices that were involved in the previous encoding steps.

Once digitally encoded, the signifying object has certain affordances. It is important to note that the exact affordances of digital objects are intertwined with their carriers; how we can interact with a digital object depends on the hardware and software into which it is embedded. Here, I will briefly touch upon the main affordances of digital objects: their flexibility, mobility, reproduction, and retention.

First of all, encoding information into digital objects affords a certain flexibility with regard to the content. Digital signifying objects can be changed relatively easily; by flipping some bits, words in files and pixels in images can be changed without leaving crossed out blotches, text can be added and deleted at any point in a document, etc. The flexibility and possibility of continuous change and addition means that the process of encoding can be potentially ongoing. Additionally, the

[2] 'Digital' is often opposed to 'analogue'. However, both are 'modes of presentation' of information and making a concrete distinction between the two is not always possible (cf. Floridi, 2009). Because the Web is without doubt a digital technology, I find it unnecessary to discuss in detail the differences (and similarities) between analogue and digital information technology. Instead, I will focus on the affordances of the digitisation of information. In this context, I take digitisation to entail the encoding of an informational unit into a discrete set of binary values (bits).

binary nature of the digital object also gives it a certain fragility: flipping the wrong bits may render it unreadable.

Secondly, the binary encoding of digitised information, allows for a precise replication of the object, without any loss of quality or quantity to the original object. The copy is generally indistinguishable from the original. With these copying affordances, digital signifying objects are infinitely expansible (Quah, 2003, p. 13-14). The copying affordances of digital objects are more than only an affordance: it is a vital part for much of the online information processing as it is a de facto necessary condition for data transmission.

Thirdly, digital objects have a peculiar 'materiality': due to their binary structure, they are not necessarily fixated to a specific location in an information carrier. They do, however, require to be stored on a physical device *somewhere*. While the carrier may be stationary, the binary character imbues the digital object with a potentially high mobility; they can easily and accurately be transported over cables and in the ether — as is done on the internet. However, in the strict sense, this is not a transportation of the object itself because this remains stored as it is on the server but the transmission of a copy. Imagine, if the original was sent, every picture on the Web would disappear from the server after the first view.

Fourthly, thanks to its binary form, large quantities of digital information can be stored on relatively small physical objects like a computer or a USB-stick. The number of digital objects that can be stored on computer chips has been substantially increased since the mid-seventies as a result of the ongoing development of digital technology (Duntemann, 1992, p. 61). For example, these days it is possible to have the complete content of a regular public library stored on a single e-reader. One of the developments that plays a fundamental role in the increasing storage capacity of computers, is the consistent decrease of the needed hardware size for storage (this became known as Kryder's law (Walter, 2005)). Another development that helped to realise the increase in storage affordances, is the exponential growth in the processing power of computer chips. The massive growth was predicted by Moore in the 1960s and became known as Moore's law (Moore, 2006). However, while there is still growth, its exponential character has stagnated and Moore's law seems to have come to its end.[3]) Nevertheless, developments of cost-effective storage maximisation are ongoing. With these storage affordances, we can often retain information indiscriminately, without ever having a need to throw digital objects away because we run out of storage space. As such, the digital storage affordances override many of the previously needed 'forgetting-by-selection' processes, in which people had to get rid of certain signifying objects in order to make room for new content (Mayer-Schönberger, 2009; Szekely, 2012). However, it is important to keep in mind that the retention of objects is not the same as the retrieval of content: the possibility to access the retained content depends on having access to equipment that can read the object's code (see section 3.3.2.3).

[3]Peter Bright, "Moore's law really is dead this time", *Ars Technica*, 2016. `https://arstechnica.com/information-technology/2016/02/moores-law-really-is-dead-this-time/`, last accessed 07-09-2018.

The above is not a complete list of the possible affordances of digital objects. By uploading objects online, these objects take on the affordances of the Web. The Web assimilates the object, and opens it up for online processing like hyperlinking and indexing by search engines (Carr, 2010). I discuss the affordances of online signifying objects and their implications in detail in this and the following chapters.

4.2.2 Perceiving digital objects

Digital signifying objects have a peculiar and multidimensional character; the binary encoded information needs to be decoded and translated into another format before it is perceptible and comprehensible to human beings. Staring at a computer chip tells us nothing about the content that it contains. The processing[4] of digital signifying objects therefore plays a fundamental and even constitutive role: the phenomenological digital object *only* exists for us through its processing. Digital signifying objects need to be processed by an output device (screen, printer, soundcard and speaker) to become available to our perception. In order to interact with digital content, we need an interface that realises the interaction between the user and the digital object.

The primary function of an interface is to allow an operator to tell the computer what to do, where to apply these instructions on, and to allow the computer to report back the results. In order for these different entities to interact, they need a shared 'language'. For this language to be manageable beyond a select group of logicians and computer scientists as operators[5], we need software that translates the bit-patterns that constitute machine-intelligible instructions and data, to bit-patterns that constitute human-intelligible instructions and information. The appearance of the digital object is therefore mediated by an interface that translates the binary data into human-intelligible representations (visuals/audio) based on a certain standard (this could be e.g., ASCII, UTF-8, JPG, PNG). As such, our interactions with the the digital entail what Ihde calls a 'double translation process'; something in the world is translated into digital code, which in turn is translated into something suitable for human perception (Ihde, 1990, p. 92). Because our experience of the digital necessarily takes place through an interface, the interface establishes what Ihde calls a 'hermeneutic relation' between us and the world (Ihde, 1990, p. 86). The interface offers users the experience of "a transformed encounter with the world via the direct experience and interpretation of the technology itself" (Rosenberger & Verbeek, 2015, p. 17). In this, the interface is more than a surface of interaction; it is also the environment and the material casing in which the interaction is realised — where it 'becomes

[4] I use the term 'processing' in line with the GDPR, which defines it as "any operation or set of operations which is performed on (...) data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction" (art. 4(2) GDPR).

[5] Though technically possible, even amongst logicians and computer scientists operating a computer by reading and writing raw bit-patterns would nowadays generally be considered unworkable.

real'. By setting the conditions for experiencing the digital world and interacting with it, the interface impresses its own intentionality on the interaction by offering users certain options, while disabling others. For example, interacting with a computer through a terminal offers a different user experience than interacting with it through a graphical user interface (GUI). In the terminal, the user needs a relatively large amount of know-how, but she is free to give a wide range of commands on the spot. Contrarily, a GUI is easier to operate on a basic level, but by requiring the user to move in certain preset trajectories by clicking on icons, it leaves the user with less freedom to operate the machine.

In the case of the Web, the interface generally consists of a device with an internet connection and a Web browser or and/or particular Web or mobile application, like a social media application. The devices that interface our interactions with the Web, come in a great variety of shapes, many of which are mobile, even pocket-size like smartphones. Combined with wireless internet, users can access the Web anywhere at any time with such a mobile device — thereby intensifying the presence and availability of the online information for users. Especially smartphones have a considerable effect on the user's relation to the online world; due to their small size and light weight they can easily be carried around and are always ready at hand to interface between the user and the Web. The developments on the level of mobile internet devices led to the constitution of a ubiquitous milieu that envelops us all (Hui, 2013, p. 52). As such, mobile devices allow users "to maintain a 'symbolic proximity' with family, friends and colleagues, whereby it promotes a sense of 'presence while absent'" (White & White, 2005). One of the results is that, for example, work-communication often enters the private time and sphere (Derks & Bakker, 2014, p. 411-412).

On the device, software in the form of a browser, generally with a GUI, mediates the relation between the user and the code that constitutes the online digital object. The browser allows users to 'surf' the Web. In its interfacing between users and online content, the browser not merely translates, but also reveals and conceals digital affordances as it offers certain options to the user. As such, the layout of the interface can foster certain actions while suppressing others, and establish certain norms for Web use (Stanfill, 2015, p. 1060).

Some of these browser functions that are relevant to mention in light of this study, are the copy functions, the search function, and the following of a hyperlink by means of a 'click'. I will discuss hyperlinks in detail in section 4.4.2. Here, I will briefly touch upon the copy and search functions.

Many browsers demonstrate a paradoxical relation to the copying affordances of digital objects. On the one hand, they make these affordances highly visible by offering users easy to execute copying functions. Images, hyperlinks and the like can generally be copied and saved on a location of the user's preference by right clicking on the object and selecting one of the copy styles, e.g., "Copy Image" or "Copy Image Location". In some cases, websites have tried to disable the browsers right-click options. However, users can try to work around this by, for example, accessing the content through the source code and copy it from there, or with the use of the printscreen function offered by interface of their device. Additionally,

some browsers themselves offer the option to disallow the disabling of right-click. On the other hand, web browsers hide and delete a significant amount of the copying that they perform in the line of regular web use. As discussed in the previous section, Web access entails the transmission of a copy. Many of these copies are automatically erased by the Web browser, without offering the user any options to retain the content. For example, when a user is watching a streaming service, the Web browser shows the content of a file to a user while still downloading it and automatically deletes the file after use. When the user wants to process a copy of this file to use outside of the Web browser, she will have difficulty getting her hands on the file. Web browsers thus promote user control with regard to copies in some contexts, while they hamper it in others. It lies outside the scope of this study to fully trace the pro- and contra-conditions with regard to the copying of online content, but some of it is likely to be motivated by copyright regulation.

The search-function is used to immediately access a search engine website with a query (search engines themselves will be discussed in chapter 6). This function is strongly represented in most GUI-browsers; it is often present as user input bar at the top corner of the browser emphasised with a magnifying glass as icon (see figure 4.1). What is more, in many browsers (e.g., Safari, Firefox, Chrome), the address bar even is the search bar, providing the function a highly prominent use-position. In chapter 6, I will discuss in detail the impact of search on the relation between users and the access of online content, so for now I will leave this aside.

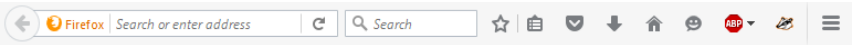


Figure 4.1: A navigation bar (Firefox Web browser with Adblocker and Privacy Badger)

By offering and highlighting certain functions, "the interface imposes its own logic on media" (Manovich, 2001, p. 76). With copy, search, and hyperlink functions, the Web expresses an intentionality towards information access (following of hyperlinks and search) as well as collection (copy, save). The browser offers users these functions, while it allows them to focus on the content they access. The workings of the browser itself are something in the background, generally hidden from the user's view. By allowing relatively intuitive operation and concealing the source code, GUI's smoothen the experience of the user navigating the Web and reduce the feeling of being mediated (Galloway, 2004, p. 65-67).[6] The smoothness of the experience is further supported by the speed with which the browser presents the content to the user; Web access should be as quick as possible and it should be avoided that users need to wait (Fielding *et al.*, 1999, p. 47).

In conclusion, as a necessary medium between online signifying objects and users, the browser expresses a significant intentionality in its mediation between the user and the online world.

[6] Although many browsers do offer the option to show the source code of a web page on request.

4.3 Production of online content

With the characteristics of online signifying objects clarified, it is now time to look at how the Web itself affects the presence of online personal information. This brings me to the first field of impact that I will research: the production of online personal information. The creation of online content is important for the informational persona, because it is the very fabric that constitutes it. In this section, I will examine the production of online personal content in a three-step inquiry. First I take a look at the 'how': how is online content produced and what are the means of production? Secondly, I will discuss the 'who' that can encode content, and how the means of production affects this 'who'. Next, I examine how the combination of the 'how' and the 'who' affects the 'what' that is produced online. To conclude this section, I will summarise the main points and their impact.

4.3.1 How: means of production

The Web allows the encoding of anything that can be digitised (e.g., text, images sound). Due to the affordances of the Web and interfacing devices, users can publish online at virtually any time and any location. In this respect smartphones again play a crucial role; with the means of production consistently in their action radius, users can encode content at any time and even update it to the latest state of affairs in real time. Moreover, with their recording options, smartphones and the like highly affect the format in which information is encoded. The combination of Web access and camera and sound recording options in these devices, make uploading of visual and audio content far easier than was possible before. Given the popularity of these devices[7], their impact on the production of online information is significant. Because their online encoding affordances are often intertwined with social media applications, I will discuss some particular aspects of their impact on Web content in more detail in chapter 5.

The effect of online encoding is generally almost immediate; the signifying object appears online with a single action or click in which the user confirms that she has finished the encoding. As such, the Web affords a certain contemporaneity, which in turn allows for a quick back-and-forth information exchange. This contemporaneity can even be bolstered with the use of automatic updating (a set of online technologies called 'AJAX' are used for this — this will be discussed in section 5.3), which does not require the user to explicitly 'refresh' the page in order to view the latest update.

Next to the ubiquitous and consistently available encoding options, over time the encoding itself has also become easier. While initially users needed to know something about programming in Hypertext Markup Language (HTML) in order to publish content online, these requirements lowered with the development of

[7] In Europe almost 70% of the population used a smartphone to access the internet in 2018. See Eurostat, "Individuals - mobile internet access" `https://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tin00083`, last accessed 09-05-2019.

applications like the Web Content Management System (WCMS, but more often used as just 'CMS'). A CMS is "a computer program that allows publishing, editing and modifying content on a Web site as well as maintenance from a central interface" (Sharma & Kurhekar, 2013, p. 258). This interface is generally a front end GUI, like the ones offered by Wordpress.org. The user does not need know how to program in order to publish web pages with applications like CMS: she can simply create a new web page with a single click (Sharma & Kurhekar, 2013, p. 258). The publishing enabled by such applications is therefore referred to as 'push-button publishing' (see e.g., Blood, 2004; Oravec, 2002). Push-button publishing often overlaps with "what you see is what you get" (WYSIWYG) editors. These editors allow agents to view and work on the content in a one-on-one visual representation of what the content will look like when it is opened in a browser. Combined, these applications turned online publishing into something that every common user could do. Heath and Motta state:

> *applications and services have enabled non-specialist users to contribute to the Web on a scale that, whilst in line with the original vision of a read-write Web, was previously unimaginable. This has been achieved by providing simple, well-structured Web forms through which users can, for example, tag photos of bookmarks, edit wiki entries, or write blog posts, using just their Web browser* (Heath & Motta, 2008, p. 78).

With the development of such applications, the skills, effort and time needed for online publishing is thus significantly reduced. By simplifying and speeding up the online publishing process, online push-button-publishing applications allow users to increase their online information production. However, the use of these applications affects the information that is produced. Generally, these applications are designed to produce similar formatted objects, like standardised blogs or web page layouts. In a certain sense, they are thus machinery for the mass production of online content.[8] As such, the production of content in the online tertiary memory is 'industrialised' (see e.g. Stiegler, 2009; Kinsley, 2015). By mobilising users in a production process for a standardised encoding of the tertiary memory, these applications articulate a certain intentionality in the creation of (a part of) the online content (this will be further discussed in chapter 5).

4.3.2 Who: shifts in the publishing monopoly

The development of publishing applications significantly lowered the skills needed to publish online, and thereby opened the door for a wide range of potential publishers: *everyone* that meets the material requirements for online encoding

[8] One can even wonder whether these applications lead to an alienation between the writer and the signifying object that she produces. I have not found any confirmation or negation of this in scientific literature, so I will leave this topic to explore in the future. What I did find, is that various elements of push-button publication can contribute to a certain distantiation, which I will discuss in section 4.3.3.

and who can work with a push-button application, can publish information on the Web. There is no over-arching agency that controls who can publish on the Web — there are however states, organisations and individuals that can try to fight specific content on a legal basis.

What is more, by allowing people to directly publish online, the user becomes "author and publisher in one" (Lessig, 2006, p. 18). Users can encode publicly accessible signifying objects about themselves and others directly online without having to receive any form of approval of a publishing agency, nor possible others that they publish about. Empowering users to publish anything they want, the Web freed them from traditional media agencies and boosted their autonomy to express themselves online. With this, the Web generated a new publishing playing field and altered the constitution of our informational landscape in its wake. As a result, traditional media like newspaper companies, libraries, archives and broadcast agencies, found themselves confronted with competitive information flows on the Web (Feenberg, 2010, p. 57). Feenberg states: "The Internet has broken the near-monopoly of the business- and government-dominated official press and television networks by enabling activists to organise and to speak directly to millions of Internet users" (Feenberg, 2010, p. 55).[9] In an attempt to safeguard their position as public information source, many traditional information agencies chose to become present on the Web. This required them to decide how to shape their online presence, i.e. which information sources they make available online and how. Often, this entailed the transformation of physical signifying objects, like books and newspapers, into a digital objects.[10] An example of this is the newspaper agency 'La Vanguardia', the originator of the content in the *Google Spain* case[11], that has decided to work on digitising their complete archive by scanning in the old content and making it available online as PDF-files.[12]

Furthermore, the affordances of online publishing gave rise to new types of publishing agents. A notable example is Wikipedia, which is an online encyclopedia that is published by an open group of cooperating agents (Raffl *et al.*, 2011, p. 608). Everyone can add or modify the content of the encyclopedia entries. The final shape of these entries is based on a consensus between the editing users (van Dijck, 2013, p. 133). The effect of this pivotal role of consensus, is that if enough people believe something is true, it becomes the truth on Wikipedia (van Dijck, 2013, p. 143). Additionally, we can see the rise of automated publishers on the Web. Examples are surveillance cameras that are live streaming their footage[13]

[9]Though currently the information flow on the Web seems to be dominated again by big players — albeit not the traditional information agencies. I will discuss this in the upcoming chapters.

[10]It is not in all cases clear how offline information flows can be translated into an online counterpart, especially in the case of libraries this can be challenging. Some libraries have therefore been experimenting with how to make content accessible on the Web (see e.g., John, 1996). Moreover, traditional media like libraries often want to retain their old information flows (Kelly, 2007, p. 78).

[11]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[12]http://www.lavanguardia.com/hemeroteca, last accessed 20-08-2017.

[13]See e.g., `http://www.opentopia.com/index.php`, last accessed 02-04-2015.

and bots that manage, edit and add content (van Dijck, 2013, p. 137-139).

Lastly, despite the open publishing structure of the Web, it is important to note that not everyone has equal access to the Web. Certain groups of the population have limited to no Web access, like the elderly, who find the use of internet complex to master (cf. Kiel, 2005; Eastman & Iyer, 2005). Such an inequality of access to the Web leads to an inequality in contribution to the Web's content (Baker & Potts, 2013, p. 187).

4.3.3 What: diverse personal information

As discussed in section 3.3.2.1, the process of encoding is a hybrid action in which human and technological intentionality are intertwined. The affordances of the Web thus also affect the 'what' that is encoded. The manner in which the Web affects the information that is encoded online, is intertwined with 'how' and the 'who' of the encoding that is discussed in the previous two subsections.

First of all, by affecting the 'who' that can publish, the Web also affects the information that is materialised. Agents can publish about whatever they like without pre-publication interference. The Web is thus used for entertainment, education, leisure and work (cf. Ferguson & Perse, 2000). The result is that Web is a rather eclectic collection of information, which contains anything from factual science to erotic fantasy stories about Cthulhu, information on world-changing events to trivialities in the lives of common people or cats. Given the diverse originators and content of online objects, we see in the Web therefore a convergence of different knowledge realms (Hui, 2016, p. 316). Here, the private, the public, the professional, all collide, combine and merge. This convergence includes all sorts of personal information, ranging from private interests, personal experiences and opinions to complete family histories and professional information.

The encoding of personal information by online users is further stimulated by the interpersonal and communicative nature of the Web (Downes, 1999). As users can encode content about themselves and others directly online without having to receive any form of approval of a publishing agency or the like, nor possible others that they publish about, they may easily reveal information unwanted by referents, even if the publisher's intentions are well-meant. An example of this is the Lindqvist case. In this case, a Swedish citizen wrote on her home page about other volunteers she was working with at a parish church.[14] On the page, she described the volunteers, sometimes their families, listed phone numbers, and also wrote about one of the volunteers working part-time due to a foot injury. This case is an example that shows that by allowing everyone to publish online, the Web affords amateur publishers to encode a significant amount of personal information in the online realm.

Secondly, the manner in which mediation style of the Web affects the user's experience of the encoding process, in turn affects the content she encodes. Due to the typical and necessary interfaced character of online information, the Web

[14] CJEU, 06-11-2003, C-101/01, ECLI:EU:C:2003:596 (*Criminal proceedings against Bodil Lindqvist*).

tends to give rise to a certain distantiation between the offline author and the online thought that she decides to materialise. With its online encoding, the signifying object becomes something 'out there', an autonomous object that can only be realised in the interface — something that is different, and distanced from the physically embodied offline encoder: many online objects have no *necessary* direct attachment to a particular offline expresser (although, over time there is an increasingly stronger connection between the online and the offline, I will discuss this in the subsequent chapters). Online, the expresser can remain (relatively) anonymous. This relative detachment between the offline expresser and her online expressions gave rise to the now famous phrase accompanying Steiner's cartoon of a dog sitting on a chair at a desktop: "On the internet, nobody knows you're a dog".[15] Online, people engage with others from a certain 'digital self-embodiment', like an avatar, that can highly differ from their physical embodiment. Users are likely to pick-up cues from this 'digital embodiment' and act in correspondence with their perception of this digital identity (Yee & Bailenson, 2009, p. 2006). Additionally, if the online disclosure of information cannot be attributed by others to the individual's offline embodied self, people are less fearful of disclosing information that may harm their self-presentation or be used against their interests (Ma *et al.*, 2016). This relative freedom of the presence of the physical self can work in various ways; it can be used by the encoder to express herself in ways that she may feel are a truer reflection of her self, but which she may not feel free to express in offline settings (e.g. a gay woman who does not feel like she can come out in her offline social environment), but it can also be used to express a virtual identity which the expresser feels is not reflective of her self (e.g. an office clerk pretending to be an adventurous knight for fun).

Additionally, the mutual invisibility, possibly even anonymity, also affects the view of the encoder on the audience for whom she encodes the content. For the encoder, this invisibility of the physical other results in the reduction of the other's presence (Berger, 2013, p. 294). Moreover, it reduces cues of authority and status, giving the encoder the feeling of a peer-to-peer-relationship (Suler, 2004, p. 324). The invisible 'other' is constructed — at least partially — in the user's internal representation system where the other becomes an 'introprojected character' (Suler, 2004, p. 323). The sense of distance is further strengthened by the fact that the Web affords asynchronous interactions between users: users can express themselves without having to deal immediately with the reaction of others (Suler, 2004, p. 322). Instead, they can suspend taking account of the reactions of others, or even decide to never look back.

With this sum of affordances that can generate a distantiation between the encoder, her content and her audience, the Web tends to lift some of the restrictive feelings that people generally experience when they express themselves in face to

[15]The cartoon by Peter Steiner was originally published in July 1993 in the *New Yorker*. See e.g., `https://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb_blog.html?noredirect=on&utm_term=.9c306830e967`, last accessed 09-05-2019.

face settings. This is what Suler calls the 'disinhibition effect' (Suler, 2004).[16] The consequence of this effect is that, online, people are inclined to encode information on controversial topics or display more extreme expressive acts (Bar-Tura, 2010, p. 237). Moreover, in the case of anonymous communication, online self-disclosure serves little social value (as people are not identifiable, they do not build up durable ties with their audience) and people are therefore more likely to disclose negative content (Ma *et al.*, 2016). As such, the Web can give rise to the encoding of relatively unconventional and extreme personal information. Adding to this, is that researchers found that over time there has been an increase in the willingness of people to participate in exhibitionism and voyeurism (Dholakia & Zwick, 2001, p. 3). Part of the motivation for such exhibitionism, it that it can empower the referent: she refuses to be humble and instead reveal who she is to the world (Koskela, 2004).

Thirdly, the industrialisation of encoding affects the content of what is encoded. For example, the introduction of push-button publishing affected the content of blogs. Initially blogs were web pages on which someone logs links to other web pages that she finds interesting. However, under the influence of push-button publication, the blogs became something that more resembles a diary (Blood, 2004, p. 54). Blood states: "Blogger was so simple that many of them [bloggers] began posting linkless entries about whatever came to mind. Walking to work. Last night's party. Lunch." (Blood, 2004, p. 54). We can see this taken up a notch on Twitter[17], which with its easy push-button-publishing and publishing limit of 280 characters, led users to 'micro blog'. The availability of push-button-publishing software thus leads not only to an increase in the information that is present on the Web (cf. Heath & Motta, 2008), but also to a shift in the kind of content that is encoded (Blood, 2004, p. 54).

Fourthly, the devices used to encode online content play an important role personal information that becomes available online. Desktops, laptops, smartphones and tablets have diverse affordances and promote different kinds of use. In many cases, encoding long texts will be far easier on a laptop or desktop, while uploading a photo that just has been taken is easier from a smartphone. Given that we carry smartphones around in our pocket and they allow us to encode personal information in mere seconds at any given time, or in any given state of the encoder (e.g., drunk), they are likely to have a significant impact on the content that is encoded online. Especially the camera in smartphones is a relevant encoding function. Sarachan states:

> *The compactness of the newest devices eliminates the decision to be a photographer on a given day. One no longer has to make the choice to take a camera to the zoo of an uncle's wedding; a camera always sits in one's pocket because it's a function of some other object. Carrying a camera has become as ubiquitous as wearing a watch used to be, before the cell phone became many people's timekeeper of choice* (Sarachan, 2010, p. 54).

[16] The extent of the disinhibition effect will differ per individual (Suler, 2004, p. 324).
[17] `https://twitter.com/`, last accessed 29-10-2019.

Once a photo is made, it can be uploaded online in mere seconds. The popular use of smartphones, contributed to an increase in photographic content on the Web (see e.g. van House, 2011). Moreover, the speed of uploading photographs with a smartphone is not only convenient, but it can also easily give rise to spur-of-the-moment actions.[18]

A last notable example of the impact of digital technology on the 'what' that is encoded online, is Google Street View.[19] Google Street View allows users to click on a map on any public road and access a recorded panoramic street view from the selected area.[20] The existence of this content is highly dependent on the affordances of digital information; the capturing and processing of this magnitude of content would be an impossible task with analogue technology. The implications for the informational persona of such new information sources can be significant, because it gives rise to personal information in a new context, or even to new personal information altogether. In the case of Google Street View, while not intended by the designers, pictorial personal signifying objects that frame individuals in a certain location, often between others, are uploaded to the Web. Despite the policy to blur faces (see image 4.3.3), individuals can be recognised — especially by those who know them.[21]

4.3.4 The how, the who and the what of signifying objects

In this section, I have discussed that the affordances of the Web can have significant implications for the 'how', the 'who', and the 'what' of the publication of signifying objects. By affecting the encoding process, the Web expresses a certain intentionality in the creation of online signifying objects. However, the actual encoding is the expression of a hybrid intentionality of the Web and the user together. In this, the Web allows publishing by a wider range of publishers than traditional media, reduces restrictions, and is able to accommodate an increasing amount of content. The open, interpersonal, and communicative character of the Web accommodates the encoding of personal signifying objects like experiences, opinions, and photos, by anyone. This also covers the publication of information by individuals about others — often (although not always) published without malicious intent, or even accidental. Moreover, because the Web's mediation necessarily takes shape in an interfaced manner, there is likely to arise a certain distantiation between the encoder and the content, even to the degree that the

[18] Though smartphones are somewhat paradoxical recording devices. Where on the one hand they invite spur-of-the-moment publications of recorded information by the user, they also often have a disciplinary effect on the surroundings on the user as they experience the presence of the mobile phone camera as a form of surveillance (Timan & Oudshoorn, 2012).

[19] The potential implications of Google Street View are a research on its own. Due to time and scope constraints, I will have to leave this aside for the future. However, I found it important to at least briefly mention that the Web also has given rise to new kinds of information sources like this visual map of a major part of the world.

[20] `https://www.google.com/streetview/understand/`, last accessed 29-06-2017.

[21] `https://www.google.com/streetview/privacy/\#service-use`, last accessed 29-06-2017.

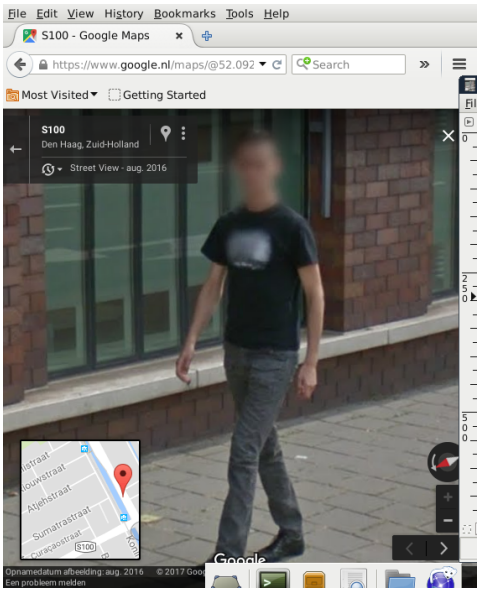


Figure 4.2: Snapshot of a passerby in Google Street View

encoder can feel free to encode more extreme content. In sum, the mediation of the Web allows for the encoding of an increasing collection of personal information, while its content can easily become highly personal, abundant and quick-and-dirty (both literally and metaphorically).

If we combine the how, the who and the what, we can distinguish between several ways in which personal signifying objects are created online. We can identify three potential encoding agents: the referent herself, another human agent, and an automated other. These three agents can encode personal information of the referent intentionally or accidentally. In the table below, I have listed the various combinations with examples (please note that some of the examples are specific for technologies that will be discussed in the upcoming chapters).

	intentional	**accidental**
referent	a 'selfie'	a reflection of the photographer in the mirror on a photo of a cabinet for sale
human other	a blog complaining about the referent's behaviour	a photo of two friends, where an unknown other passes by in the background
automated other	a social media publication reporting that the referent "likes this page"	the display of a passer by in Google Street View

The who that encodes personal information, as well as the intentions that lead to the encoding, play a pivotal role in the selection of the content that is encoded online. As such, one or more of these agents is necessarily responsible for encoding the signifying object that may cause problems. Moreover, as these agents are the driving force behind the publications, they are stakeholders in the balance of interests that comes with the application of art. 17 GDPR, which I will discuss in chapters 8 and 9. However, unless made explicitly visible, the who as well as the intentions of the encoder matter little for *the manner* in which a signifying object represents a particular referent to Web users: once encoded, personal signifying objects have a certain semantic autonomy that does not necessarily coincide with the intentions of their author (see section 3.3.2.1).

4.4 Presence

In this section, I examine how the Web affects the presence of the information that it contains. For this, I will first focus on the presence of the Web itself. Next, I investigate the manner in which information is integrated in the Web and how this affects the presentation and the context of the content. Lastly, I discuss how online content fares over time.

4.4.1 Proximity of the proxy

The Web is always 'on' and accessible from almost everywhere as long as we have access to a device that can interface our interactions with the online realm. Especially smartphones imbue the Web with a strong presence, because users generally carry theirs on them for indoor- and outdoor activities (Wang *et al.*, 2016, p. 59). With the help of these devices, the Web is almost always within our action radius. Users unlock their smartphones on average eighty times a day.[22] People are therefore often in an almost permanent state of connection, increasing the chance of a high integration of the human cognitive system with the online information flow (Lemmens, 2014, p. 2). The active bidirectional character contributes to this by constantly inviting users to interact or respond to digital objects, irrespective of where they are (Feenberg, 2010, p. 54) — thereby giving rise to a flow of constant updating. This behaviour is generally reinforced by informational rewards that the updates and reactions provide (Oulasvirta *et al.*, 2012).

With this continuous availability, the Web became part of our informational routines and gives rise to a 'hyperconnectivity' of its users (cf. Quan-Haase & Wellman, 2005; Floridi, 2015). Quan-Haase and Wellman define 'hyperconnectivity' as "[t]he instant availability of people for communication anywhere and anytime" (Quan-Haase & Wellman, 2005, p. 251). This hyperconnectivity is fostered by the increasing societal implementation of the Web as the main sphere of interaction and organisation. The hyperconnectivity of users is even so strong,

[22] Ben Bajarin, "Apple's Penchant for Consumer Security", *Tech.pinions*, 2017. `https://techpinions.com/apples-penchant-for-consumer-security/45122/`, last accessed 16-09-2017.

that many people can feel obligated to actively maintain online social interactions of their daily lives while on holiday (Wang *et al.*, 2016, p. 59). Much of Western life in the form of communication and interactions with others or in society in general, seems to be inherently intertwined with online interactions. As the Web became an important medium of the social interaction, the time we spend online has significantly increased. Also, many of our physical actions come into being as a result from an interplay between the offline and the online. For example, to catch a train, many people will first consult an online travel application in order to learn about the times, locations and potential delays, before physically going to the station. Without access to the Web, individuals are excluded from a part of societal life, such as job vacancies that are only published online. The 'real life' of individuals therefore takes place off- as well as online — this has been captured by inter alia Hildebrandt and Floridi in the concept of 'onli*f*e' (Hildebrandt, 2015; Floridi, 2015, p. 42).

Because the Web is consistently at hand, highly present, and used for many aspects of societal life, many users will turn to it first if they have a need for information. Due to the relatively effortless access, users even need little motivation to venture online; the wish to 'kill some time' is sufficient (Oulasvirta *et al.*, 2012, p. 113). The result of this low effort-threshold is that information sources that in their physical form received moderate attention, now experience an increase in their audiences. We can see this in for example the online use of archives: "Now, however, millions of people who cannot or do not want to go to the archives are accessing them in digital form" (Stallybrass, 2007, p. 1581). As such, online information — including historical information sources like archives — are "being appropriated and transformed into part of our daily material lives" (Manoff, 2010, p. 392). Personal information that was stored and left gathering dust in physical buildings, gained an audience by being placed in everybody's reach with a few clicks. The easy access and immediate presence even bestows the Web with a role as "primary form of external or transactive memory, where information is stored collectively outside ourselves" (Sparrow *et al.*, 2011, p. 776).[23] In this role, the Web affects how we think and what we remember, as it shifts our inclination from remembering content to remembering how and where to find it (Sparrow *et al.*, 2011, p. 778).

The dominant role of the Web in people's lives imbues online personal information with a strong presence. As the user oscillates between online and offline information flows, online personal information is in a state of consistent semipresence, always ready at hand with a click, a search or a command. Online information renders people present in a 'de-spatialised simultaneity', a situation where "distant others could be rendered visible in virtually the same time-frame,

[23]The notion of 'transactive memory' originally derived from Wegner's research on the processing and structuring of information within a group of people (cf. Wegner, 1987). People can use each other as external information storages. By exchanging information, their personal memory becomes transactive (Wegner, 1987, p. 186). However, the transactive memory system is not limited to human-to-human interaction, but can also be mediated by technologies (Sparrow *et al.*, 2011, p. 778).

(...) even though they did not share the same spatial locale as the individuals to whom they were visible" (Thompson, 2005, p. 37). As such, the referent maintains a virtual presence that willingly or unwillingly serves as her online proxy. This virtual presence in the form or particular objects can affect her interactions in the here and now by always being an auxiliary information source for users next to the physical setting. The online information from one context can thus always 'hoover' over the referent in other settings and influence how users view her. Online personal information can therefore have a deep impact on people's lives. Moreover, because the Web is used for private, professional, and public affairs, we see a merging of different contexts that also seeps into the offline world and disperses traditional spatial and temporal boundaries that demarcate various private and public realms. With such spill-overs of different contexts into each other, like work and leisure contexts, it can become difficult for people to maintain distinct roles in different settings. This is complicated further by the manner in which the Web assimilates personal information, which I will discuss in the next subsection.

4.4.2 Integration in the network

As the opening quote of this chapter already pointed out: the assimilation of personal information by the Web has a tremendous impact on the affordances of this information, as well as on the manner in which users experience and use it. In this subsection, I will examine what it means for personal information to be integrated in the Web.

Online information is embedded in web pages. However, in order to access this information, a user needs to have the location of the web page. The location of the information is identified by Uniform Resource Locators (URLs, e.g., www.existentialcomics.com). The URL identifies information as a resource at a particular network location (Berners-Lee *et al.*, 2005, p. 7). URLs have a scheme-specific syntax, starting with a reference to the protocol that locates the resource, the Hypertext Transfer Protocol (HTTP). This forms the familiar "http://". The location names are given shape in a for humans reproducible form in the Domain Name System (DNS).[24] The DNS allows the use of keywords and provides a general hierarchic naming-system. By moving from the most general last extension ('.com', '.nl', '.de', etc.) to more specific parts of the domain-name, the DNS shows some familiarities to a phone book. However, in this version the name is integrated with the number and far less ordered; agents are relatively free in their choice of the domain name, particularly since top-level domains have opened up to a far broader spectrum (see the expansion of generic top level domains, gTLDs[25]).

In their human-friendly form, URLs often already reveal or suggest something about the information that can be found at the specific location. However, the content embodied in the URL can easily be a mismatch with the content of the actual

[24]Other URL-schemes exist. One of the typical reasons for designing and using another URL-scheme is because agents do not want to conform to the DNS (Berners-Lee, 2011).

[25]See ICANN press release, `https://www.icann.org/en/system/files/press-materials/release-19jan12-en.pdf`, last accessed 16-05-2019.

resource. Take for example the following URL: https://joop.bnnvara.nl/nieuws/partij-voor-de-dieren-helpt-knetterrechts-college-aan-meerderheid-in-limburg[26]. Roughly translated, the last two parts of this URL state: "news: Party for the Animals helps extreme right to a majority in Limburg's board of governors". Contrarily to what this URL suggests, the resource to which the URL refers, actually states: "Partij voor de Dieren helpt 'knetterrechts' college *toch niet* aan meerderheid in Limburg" [my emphasis], which means that the Party for the Animals does *not* help extreme right to a majority. Users who encounter the URL and who do not open the actual resource, are therefore likely to draw the wrong conclusions about the Party of the Animals. Moreover, URL names can even be intentionally used to lure in users to confront them with unexpected content. A notorious example of such a 'bait-and-switch' website is www.lemonparty.org, which is not about lemons.[27]

When a user arrives at the web page location, she generally encounters a signifying object that is displayed between other signifying objects. Because the surroundings of a signifying object are necessarily taken in by the user, this context affects her experience of the information (Carr, 2012, p. 485-486). The objects therefore shape each other's context and affect how they all are interpreted. However, an online signifying object is not just embedded on a single web page, but it becomes part of the Web itself, which is, indeed, a *web* of information. The Web affords the realisation of interactive connections between signifying objects by means of 'hypertext'. Hypertext is a form of writing that allows the creation of interconnections that are automated upon request and "go somewhere, do something, 'perform' or expand" (Barnet, 2013, p. 6). Hypertext is code that implements interactive links with and to other resources, and allows for a dynamic treatment of the content. The active character of the hypertext is realised by hyperlinks; these contain URLs to other resources and allow immediate access. A hyperlink is a relationship between the link's starting point, its head anchor, and the location where the link goes to, its tail anchor (Berners-Lee & Connolly, 1995, p. 37). The anchors are not connected to a specific signifying object, but to a specific location in the network (Berners-Lee & Connolly, 1995, p. 38). With a single click on a hyperlink, a user is immediately directed to the hyperlink's target location. As such, hyperlinks can bridge content between two different web pages, thereby creating a relation between signifying objects in different contexts and sources. However, this action is one-directional: the location to which a link directs the user, does not automatically refer the reader back to the starting location. This unidirectionality allows for the existence of what is known as 'link rot', where a link points to content that has been removed. Hyperlinks can have various formats; they can for example be displayed as text, image, or icon (Berners-Lee & Connolly, 1995, p. 30). Also, hyperlinks can be embedded. In this case the user is not directed to the target location, but instead, the content at the target location is displayed within the context of the head anchor page.

[26]Last accessed 30-01-2019.

[27]Know Your Meme, "Lemon Party". `https://knowyourmeme.com/memes/lemon-party`, last accessed 21-08-2017.

The hyperlinks provide for a link structure in the Web's information collection (Page *et al.*, 1999, p. 1). As such, hyperlinks deeply impact the affordances and appearance of online information. They play an essential role in the Web's interface and structure, and are frequently used (Obendorf & Weinreich, 2003, p. 736). The hyperlink's unidirectionality allows for hyperlinking to resources without the need of consent or confirmation of the target-resource. This means that resources and their signifying objects can be represented in a context or associated with certain other resources that may go against the intention or the wish of the original publisher. Not only is there no need for consent in order to establish a hyperlink, but the controller of the linked-to content may not even be aware that the hyperlink exists.[28]

To signal users that they allow a certain action, hyperlinks are often designed to visually stand out and attract attention (Obendorf & Weinreich, 2003, p. 739). The hyperlinks can be underlined, displayed in bold font, or coloured in a way that sets them aside from rest of the content. This visible salience increases the likelihood that a user will use the hyperlink (cf. de Ridder, 2002). By attracting attention, hyperlinks increase the presence of the references that they point to. Moreover, on a semiotic level, hyperlinks signify a *suggestion* for retrieval. This shows the socio-technological nature of the hyperlink; it is in a way a digital finger pointing in a certain direction — and conveniently, also paves the road to take us there. As such, hyperlinks signal meaning to the particular associative relation that they establish between the content of the head anchor and that of the tail anchor. Their signalling character becomes a seductive detail in the source text and invites the user to follow the link. This 'seductiveness' can affect a user's perception and comprehension of the informational context of the hyperlink: "Readers focus on a seductive detail and remember it, sometimes at the expense of the target point of a text. They may also misinterpret the text under the influence of seductive details" (Wei *et al.*, 2005, p. 435). Hyperlinks thus increase the qualitative and quantitative proximity of their target by attracting attention, signifying meaning, establishing an association, increasing the access points and by accelerating the access process. The consequence is that hyperlinks can enhance or reduce the presence of a particular reference, affect its framing and lead to spill overs of one informational context into another.

Lastly, the format of the content plays a role in the manner in which the online informational persona is present. The improvement of internet-connections and the development of user devices with more processing power afforded a better and faster display of graphics. Many web page designers made use of these affordances and, over time, the Web became a more visual medium where pictorial content is abundantly used (Singer, 2009, p. 375). With this, the collection of online personal signifying objects is thus also likely to become increasingly pictorial. Many Web developers even "hold the view that the Web is, by nature, a graphical medium and therefore is the domain of the graphic designer" (Rowan *et al.*, 2000, p. 80).[29]

[28]There is no notification or registration of hyperlinks. However, a web page controller could derive from her log files that certain other web pages direct content to her web page.

[29]This is in conflict with the W3C's view of the Web, which explicitly states that the Web should

Taking this all into account, we can conclude that the assimilation of personal information by the Web has tremendous implications for the presence of references. Once online, the references embedded in signifying objects become present between other items, which mutually influence each other's interpretation by users. Moreover, the content can be moved, copied, linked to, and searched, thereby creating new contexts for the content and also possible new variations (in chapter 6, I will discuss the impact of digital search provided for by search engines, and in chapter 7, I will specifically look at the effects of the spread and multiplication of personal signifying objects on the informational persona). Especially the hyperlinks can play a significant role in this all; as a user chooses her own path way through the Web of hyperlinks (and in some applications the user will even receive personalised links, this I will discuss in more detail in chapter 6), each user will likely get a different view of the informational persona.

4.4.3 Personal information over time

Once online, personal signifying objects shape the presence of references in the manner discussed in the previous two sections. However, because as human beings are not static in time, it is important to also consider how the Web relates the presence of references *to* time, as well as how it mediates them *over* time.

Let us first have a look at the manner in which the Web relates personal references *to* time. This may be best considered an 'anti-relation', because the Web does not provide a chronological overview of the information that it contains. On the contrary, as hyperlinked web, its main structuring features consist of association, which easily crosses different temporal origins of signifying objects. Moreover, it can be difficult to discern the temporal context of online signifying objects. The online view that a user has on a particular referent, is therefore in all likelihood a-chronistic (unless she encounters a chronologically ordered personal web page, and even then it is just the single page among the rest of the Web).

This brings me to the Web's mediation of personal information *over* time. The main concern with regard to online information is that the Web would entail an 'everlasting memory' (see chapter 1). Due to the massive storage capacities of contemporary servers, Webmasters can retain online information indiscriminately, without a need for deletion to free space. As a result, the Web could potentially grow in into an ever-expanding tertiary memory that allows us to revisit everything that was ever produced online. The case of the referent that after 10 years was still confronted with the website that she made when she was in college is an example of this (see section 4.1). However, ongoing unchanged retention does not seem to be the status quo of the Web. First of all, the ongoing retention of online content requires maintenance. It takes action and effort to keep the hardware and software underlying the web pages viable, updated, and working. Secondly, personal information does not often remain available as-is on the Web. On the contrary, the Web gives rise to a dynamic informational environment where

be able to deliver various kinds of content in order to serve people with different capabilities. https://www.w3.org/TR/2004/REC-webarch-20041215/, last accessed 07-08-2019.)

information is moved, removed, updated, and edited. The online culture is one of constant updating and editing (Carr, 2010, p. 107). Last, the initially indivisible (discrete) nature of digital objects makes these objects relatively fragile; they can easily be overwritten or rendered unreadable by minor damage (Vafopoulos, 2013, p. 79). The connective character of objects on the Web adds to this fragility. Online signifying objects can become (temporarily or permanently) inaccessible due to issues with the domain name, URL or server. If a domain name owner stops paying for the domain or if the server breaks down, a page 'disappears' from the Web. In turn, hyperlinks also easily deteriorate; if a web page is moved, the hyperlinks on other pages pointing to the content's initial location 'break' and give a 404-error. While there are attempts to reduce these problems, by for example the use of permalinks or persistent uniform resource locators (PURLs), both of which aim to provide more persistent URLs, link-rot on the Web is still a regularly occurring phenomenon.

Online, we thus see that signifying objects appear, disappear and change on the Web, move from their locations, are mirrored on other servers and so on. Moreover, there is no backup device for the Web (Fuller, 2003, p. 69). While some attempts on this front have been made in projects like the 'Waybackmachine'[30], these projects only cover a small portion of the Web and tend to be spread across different devices and controllers, making them difficult to find and access (Fuller, 2003, p. 69). The Web is therefore "not a perfect archive: information gets lost, items are daily replaced or removed, content is duplicated all over the shop. There is no guarantee that the article you visited last week will still be there today" (Barnet, 2013, p. 138). However, this does not mean that the Web does not cause trouble for the presence of personal information over time. A part of the content does remain available over time due to the lift of the previously necessary forgetting-by-selection (if the required maintenance is performed) and the human impulse to preserve information (Manoff, 2010, p. 386). Moreover, if a signifying object is copied, it can always be uploaded again. With the combination of these factors that can secure the availability of information over time, an online personal reference may gain a certain persistence (van den Berg & Leenes, 2010, p. 1112).

4.5 Publics

The third element that I discuss in this chapter, is the public of online personal signifying objects. Just like the encoding, the accessing of content is a hybrid affair. In order to examine the public that results from this hybrid affair, I will split up the inquiry into three parts. For this, I will start by focusing on the two agents of this hybrid intentionality separately. First, I will focus on the impact of the Web itself as a technology on the composition of the public. Next, I will discuss the impact of users on the formation of online audiences for signifying objects. Lastly, I will combine both and examine in more detail how the public is formed by their hybrid intentionality.

[30] http://www.wayback.com/, last accessed 07-08-2019.

4.5.1 The impact of the Web on its publics

The Web restricts the publics of the information that it contains to its users. First of all, to be able to use the Web, people need to have access to certain devices and resources. In the design of the Web and its applications, certain preconditions are set for the hardware and software needed to access online content. Over time, Websites and their content became more complex and demanded more processing power of user devices (Hargittai, 2003, p. 259). As such, users with older devices or software may have trouble accessing content on flashy contemporary websites. Secondly, the technology places certain demands on the know-how of people. However, this necessary know-how has been heavily reduced over time (see section 4.2) and the required devices became easier to operate (e.g., tablets). Groups that were initially disconnected from the Web, like the elderly (see e.g., Kiel, 2005; Eastman & Iyer, 2005), are now increasingly more likely to find their way online (also, because with the passing of time, the generations familiar with the internet slowly become the new elderly).[31]

A substantial part of the world population meets the above mentioned conditions to a greater or lesser degree, and is an internet (and in all likelihood Web) user. Estimation is that 82,9% of the European population was an internet user by March 2019.[32] Worldwide the number of internet users is estimated at over 4,346,561,853 — which is over 50% of the world population.[33] Given the significant size of this user group, from here on I will keep my focus on those who are Web users.

The technology of the Web affords massive indiscriminate audiences. In theory, the only strongly restrictive factor is that someone needs to be a Web user. The Web imbues online objects with a high accessibility compared to spatially dispersed physical signifying objects (Nissenbaum, 2010, p. 125). Access to them is fast and cheap, while spatial distance is nullified as a demarcating factor. The accessibility is further increased by the fact that web pages are non-rival goods; this means that the consumption of the good — viewing the information — by one person, does not diminish the usefulness of and access to the good for others (Quah, 2003, p. 13). Web technology thus allows a significant number of people to retrieve the same online signifying object at the same time, without any of them preventing another person from retrieving the same content.[34] Being online accessible, therefore has a crucial impact on the public of a signifying object; the object in theory becomes instantly accessible to a public that could potentially consist of over half of the global population. However, in practice, the actual public of a particular personal signifying object will be shaped by several other factors, as I will discuss now.

[31] Of course, people may also consciously choose to refrain from Web use.

[32] www.internetworldstats.com/stats.htm, last accessed 07-04-2019.

[33] Ibid.

[34] Still, there is a limit to the maximum number of people that can view a website at exactly the same time due to the burden that this places on the internet's infrastructure (Vafopoulos, 2013, p. 84). However, this is a relatively minor limitation compared to the offline world.

4.5.2 The impact of the user

The user, as a human being, also plays an important role in the composition of the online public for a particular signifying object: she takes in an active role in accessing content. Given the low effort and costs needed to retrieve information from the Web, users will only need a minimal interest or urgency to retrieve online content (being bored is even sufficient motivation (Oulasvirta *et al.*, 2012, p. 112)). However, there are two elements that affect the likelihood that a user takes notice of a particular signifying object: her attention and her background.

A user cannot speed up the time she uses for attention, memory and/or imagination beyond a certain limit (Berardi, 2011, p. 55). Moreover, a user has only a certain amount of time available to her to surf the Web. The result is that no Web user has the time to spend attention on *all* the content of the Web; there is too much information available. The abundance of online information therefore creates a scarcity of attention (Simon, 1969, p. 40-41). This scarcity makes online attention valuable and gave rise to an 'attention economy' (cf. Goldhaber, 1997). The attention of users is valuable not only due to its scarcity, but also due to its potential consequences; once an agent has the user's attention, she can potentially steer the mind and body of that user (Goldhaber, 1997).[35] Online, we can therefore often see a battle for the attention of users going on, in which several players try to grasp the user's attention with diverse techniques like the use of moving or flashing images. The strength and the volume of the signals play an important role in this: generally, 'loud' signals will attract the most attention (Falkinger, 2007, p. 268). Eye-catching and easily digestible content like images, headlines and hyperlinks, are more likely to be noted by a user than plain textual content, and thus receive bigger audiences. This is strengthened by the fact that on the Web, users tend to quickly scan through content (Obendorf & Weinreich, 2003, p. 741).

Furthermore, the user's background will likely attract her to content that she understands and which matches her interests. At the minimum level, this will affect the likelihood that she will spend attention on a particular textual and/or pictorial object. In case of text, the presence of the content depends highly on the language written in the object and the language of a particular user. Because most users will focus on content that they can understand, they will likely ignore signifying objects in a foreign language. Pictorial signifying objects (images or video) that do not depend on the written word, are easier to interpret across various linguistic networks. Moreover, because pictorial content also tends to be easily digestible and consumable, it can grab the attention of a larger public.[36] As such, pictorial content can become a highly present personal reference (this will be discussed in more detail in chapter 7). However, because pictorial content is more easily taken in by a larger audience than a piece of text in a particular language,

[35] Attention plays an important role with regard to the position taken in by gatekeepers, as I will discuss further in chapters 5 and 6.

[36] However, pictorial content may rule out a part of the user group: the visual impaired tend to rely on text-to-voice software in order to interact with the Web. As pictorial content falls outside the scope of these programs, the "increasing use of graphics (...), is blocking out people with disability" (Ellis & Kent, 2010).

this content may be at a greater risk to be misunderstood, since images can have different meanings in different cultures. While pictures thus may be easily shared across linguistic communities, they can still suffer from interpretation errors.

4.5.3 Public composed in a hybrid intentionality

Both the Web and the user thus bring into play some relevant elements for the composition of online publics. If we combine attention, background, and worldwide easy access, we see the emergence of a public that revolves around interests and cultural background.

The user herself takes an active position in the accessing of content (e.g., click on hyperlinks, browse) and ventures on an informational journey that may be affected by attention-grabbing loud signals, but is also fuelled by the user's interests and background. The global scope of the Web allows users to find, unite and interact with others with similar interests, life-styles, problems etc. — even if it is a rare common denominator (Feenberg, 2010, p. 56). The shared interests can be as general as a shared language or so specific as a love for DIY synthesizers. As such, the networked and bidirectional character of the Web gives rise to smaller cultural networks or 'clusters' that evolve around common interests (Lovink, 2005, p. 18). By giving rise to interest-based networks, the Web entails a re-clustering of the relation to information by replacing the offline clustering that was usually based on kinship or geographical vicinity, with a clustering based on interest (Wellman, 2001, p. 13). While the clusters evolve around interest and are not necessarily connected to the offline world, users do tend to have a greater or lesser degree of overlap between their online and their offline connections (see e.g., Reich *et al.*, 2012). For example, research in social media has shown that people use online interactions to reinforce relationships that already existed in the offline world (Kim *et al.*, 2011).

On the information flow level, people tend to stay within informational cultural networks that match — at least partially — their own views, ideology and/or social demography (Nahon & Hemsley, 2013, p. 75-76). This confinement to personally matching content is often strengthened by technologically driven applications that profile the user. Based on this profile, the mediating technology envelops the user in a kind of 'filter bubble'; the user is shown the content that matches her profile, while content that does not fit the profile is filtered out (Pariser, 2011). I will discuss these mechanisms in more detail in chapter 6.

The consequence of the cultural subnetworked character of the Web is that, despite the potentially global reach of online information, personal references are generally present in particular interest networks. The presence of certain signifying objects in a particular cultural network depends on the relevance of the object for that network. Certain content will be more popular and raise more interest in specific networks than it will in others and therefore is (more often) uploaded or linked to. Given the interest-focused clustering of users, the presence of a particular reference will likely be centred in a cultural network or set of networks with which the carrying objects share certain denominators like language and/or professional

interest. With these common denominators there is a likelihood that the online audience overlaps with the past, present or future offline audience of the referent. The presence of a certain reference for a specific user will thus depend on the extent to which the carrying objects are embedded in one of the user's interest networks. However, it is important to note that because people generally have multiple interests, they tend to participate in various networks (Wellman, 2001, p. 15). The interest networks are therefore fluid. As such, personal information can easily be introduced in new networks by a user moving between cultural networks (Wellman, 2001, p. 15).

However, this is not the full story of online publics. On top of the standard Web, several technological applications can be put in place that intervene with the potential publics accessing particular content, or steer users towards particular content. Examples of restrictions to content are interventions raised by governments or industry, like the blocking of particular websites in specific countries (Leenes, 2011, p. 156). Also, web page controllers themselves can restrict the access to their content by for example placing the access to the content behind a password or a paywall. On the other hand, controllers can also affect audience access and composition by steering audiences to particular content or pushing the content towards them. An example of this is the use of feeds and search functions. Feeds and search functions can be used to connect audiences with content based on various selection criteria like the user's geographical location. These technologies are typical for the audience composition on social media websites and in search engines (see chapters 5 and 6). I will therefore leave the detailed analysis of these mechanisms for chapters 5, 6, and 7. As I sketch a general image in this chapter with regard to regular web pages, I will for now suffice with the remark that the application of such additional technologies can heavily impact the presence of particular references by restricting the audiences of certain signifying objects, or on the contrary, by pushing them towards content.

4.6 Complications of the presented persona

The Web as tertiary memory serves as a protention for its users' understanding of the world (see section 3.3). Correspondingly, personal information materialised on the Web, serves as a protention for how users perceive others, as well as themselves. What is online, who can access this, and when, is therefore important for the manner in which people are understood by others and themselves. However, the Web itself as a mediating technology affects the processes of encoding, storage and retrieval of information and thereby presses a certain technological intentionality on this information and its presence for users. In this section, I will combine the findings of this chapter and discuss how the Web mediates personal information and thereby co-constructs for users a certain view on the referent that may represent her in a problematic manner.

In the first place, the presence of the Web itself for users matters. The Web's continuous and ubiquitous accessibility (especially with mobile devices like

the smartphone), combined with its social implementation, bidirectionality and hyperconnectivity of its users, establishes the Web as highly present, prominent and pervading tertiary memory, that is constantly within the action radius of its users (see section 4.4.1). The prominent role of the Web in Western society even gives rise to expectations of a certain online presence; connectedness is assumed. The consequence is that, even personal information that already used to be public, like information that was originally physically published by traditional media, gains a different — generally stronger — presence when assimilated by the Web. However, the presence of this tertiary memory has a somewhat paradoxical character. While highly present and at the user's fingertips, the Web is at the same time necessarily mediated by an electronic device that is always between the sender and the receiver (see section 4.2). As a result, the sender and her audiences may feel distanced from each other — hidden behind an interface and removed from each other's immediate reach. The Web as a tertiary memory therefore has a high proximity to users, but always at the expense of an unavoidable distance. This distinctive presence of the Web combined with its digital affordances, affects the encoding and retrieval of personal information — which in turn also affect each other.

On the level of encoding, the Web allows users to be an author and publisher in one (and an anonymous one if they desire so), who can publish anything online at any time. Bolstered by the industrialisation of encoding in the form of push-button publishing applications (section 4.3) and combined with their hyperconnectivity, users are invited to react, publish and interact online. This easily gives rise to the publication of a vast amount of personal information online. Additionally, the distantiation generated by the interfaced interaction may give rise to the encoding of less nuanced and more extreme personal information, than face-to-face contact (see section 4.3.3). Spur-of-the-moment actions, emotional outbursts and private revelations may all find their way to the Web (I will discuss some illustrative examples of such behaviour in chapters 5 and 7). As users encode content online, they share information about themselves, but often also about others who spiked their interests or crossed their paths. This sharing of information about others may even be accidentally, like a passerby in the background of someone's holiday photo. The result is that a significant amount of personal information is encoded on the Web — either by the subjects themselves or by others. If content is encoded by others than the referent herself, these others create a particular part of her persona that may expose or display the referent in unwanted manners and undermine the control over her own self-presentation. Some of this online personal information may even be created without the referents being aware of it. Others — in an interplay with the technology — can therefore be an important cause of problems.

Meanwhile, on the level of information retrieval, the Web can pose a challenge for senders by affecting the audience composition. As discussed in section 4.5, the networked character of the Web gives rise to a form of compartmentalisation based on a mix of interest, cultural background, and the user's offline social network. In general, we can expect personal information relating to a particular referent to be more present in a certain cultural network, the more the referent or the publisher

of the content shares common denominators with the users in this network, like language, country of origin, interest, political views, etc. However, it is important to note that this compartmentalisation is highly fluid: online the content is open to an indefinite audience and due to the affordances of online information, it can quickly and easily be injected into different networks by users (this will be discussed in more detail in chapter 7). Online personal signifying objects, even those referring to the trivial and the local, can therefore potentially reach significant audiences. The open access and easy spread of personal information across multiple audiences can challenge the referent's online self-presentation when performing distinct social roles: even if referents themselves or benevolent others share information only within a critically selected cultural network, the content is open to unforeseen or unintended audiences, or may easily reach these by means of republished copies. An example of users reaching unintended audiences, is when users intend to share their holiday experiences 'live' with their friends and family in order to keep in touch, but also inadvertently end up tipping off burglars on which houses to rob (especially social media are used for this because they allow a relatively easy narrowing down of the victims, which I will get to in chapter 5).[37] Even content meant to be public may reach an audience beyond the user's expectations: a publisher will necessarily (although not always consciously) publish for a particular public and is therefore likely to have formed certain expectations with regard to the identity of her public, like its lingual, national, and cultural identity, as well as it having a particular interest in the content. Due to the Web's global scope, a publisher may easily have overseen potentially unwanted publics. The impact of the Web's mediation on the scope of the public was one of the major points of attention for the CJEU in the Lindqvist case (already briefly discussed in section 4.3.2), which revolves around the publication of information by a Swedish citizen about volunteers in a church parish.[38] Despite Lindqvist's good intentions and the fact that she did not publish the content for economic gain, the publication was considered to be an infringement of the rights of the referents to whom Lindqvist referred on her website, because their personal information was "made accessible to an indefinite number of people" (§47).

Overall, the consequence of being the referent of online signifying objects is thus that an individual can gain a tremendous visibility in the form of a compilation of one or more particular references. And this visibility is not fully under the control of the referent, or is even not under her control at all. The exact appearance of the persona depends on the available content in combination with the hybrid intentionality of the viewer as well as the Web. In this, the Web impresses its typical networked character on the formation of the persona as it is presented to users. Online, the individual is not represented by a stand alone set of signifying objects, nor by a chronological collection. Instead, the persona is shaped by a variety of objects that are woven into the bigger fabric of the Web, where they are

[37] For example, see `https://getsafe.com/how-burglars-use-social-media/`, last accessed 23-11-2019.

[38] CJEU, 06-11-2003, C-101/01, ECLI:EU:C:2003:596 (*Criminal proceedings against Bodil Lindqvist*).

embedded between other content, become open for copying, editing, hyperlinks and are incorporated into particular interest networks. Especially hyperlinks play an important role in this; they can significantly increase the salience of particular personal references by pointing towards them, while they also can establish associative relations between different objects and reveal something about the content to which they point. The hyperlink draws the user's attention to the linked-to content as it forwards her with lightning speed to the target location. As the user follows a trail of her interests through various hyperlinks, her attention per visited page curbed by her own attention span, an image of the referent emerges, that is less the result of an intentional act by a particular human author, than it is of the associative movement afforded by the medium. Identity, as portrayed by the online informational persona, is therefore one of associations and connections that are more loosely or haphazardly combined than used to be the case with offline personae: the online informational persona is one in which the past and present, the far away and the distant, are blended.

With the blending of objects, contexts, time-frames and associations, the original content and context of a personal signifying object is easily distorted. Many of the affordances of online information therefore pose a substantial risk to the contextual integrity of the informational persona (cf. Nissenbaum, 2010). The associative framing, potential distortion and decontextualisation of content, may inadvertently imbue a reference with a different meaning on the referent than was originally the case or the intention of the author, resulting in a growing risk of users misinterpreting the referent. Moreover, as objects can be edited without giving a hint to the viewer that the object has been manipulated, they may give rise to problems on the level of interpreting the authenticity of the object (Gregory & Losh, 2012). The online informational persona as viewed by a particular user, may therefore easily reflect the referent in a problematic manner.

Moreover, the impact mentioned above can be prolonged over time, because online signifying objects tend to remain stored by default if the servers are properly maintained. In these cases, erasure requires an action or an accident (see section 4.4.3). Thus while the Web is not a perfect archive and online content is volatile, online personal information may be retained ongoing. The consequence of this is that, over time, the individual can be represented by a voluminous online informational persona consisting of a vast array of signifying objects — and it can keep on growing.

Taken together, by giving rise to an informational persona that takes shape for the view of a particular user in the form of a-chronistic weave of associations, the Web constitutes a challenging environment for the construction of an informational persona that matches the referent's self-perception. Meanwhile, the impact of an online informational persona on the referents' lives can be significant, because the Web is the main communication pillar of contemporary life. By being assimilated by the Web, a problematic reference takes on the Web's affordances and is endowed with a presence which allows it to complement, negate or even overrule the offline presentation of the referent's persona. Irrespective of whether online signifying objects are true or false, they can highlight certain personal information, bring

it under the attention of Web users and into their ongoing frame of reference. The online personal information can spill into the referents' offline lives at any moment and challenge the context and presentation of their offline interactions, the different roles that they may want to play, and thwart their attempts to segregate their audiences (cf. Korenhof, 2014). The online references could even become prevailing elements in the constitution of an individual's informational persona compared to offline objects. The result is that the online persona may leave users with an impression of the referent, that is a poor or even erroneous reflection of who she is (now), or is a reflection that she disagrees with. Moreover, if the referent herself is the Web user who is confronted with the content, she may find unpleasant memories triggered or her self-view questioned in troublesome manners. The online persona can thus hinder an individual by framing her in particular predicates for both the perception of others as well as her own perception. A high presence of a certain reference may therefore to an inability to move past it and thereby hamper individuals to heal from previously experienced traumas (Holman & Silver, 1998). Even online expressions that were meant only as an expression of a virtual identity for a virtual audience, may spill into the offline world and reflect back on the referent. When this happens, the individual may find her physical self overlaid with these expressions of a virtual identity. As more and more of our interactions move into the digital realm, the online persona in many cases even replaces the person herself as the object of decision-making (cf. Clarke, 1994; Zwick & Dholakia, 2004; Roosendaal, 2009). Hence, "the virtual has real effects — either on those who live it, or on those who live with them" (Lessig, 2006, p. 20). At the very least, the impact of the Web's mediation of personal information seems to provide sufficient reason for wanting to have something like art. 17 GDPR to address certain problems.

Before delving into the question of whether art. 17 GDPR is actually equipped to be of help to counterbalance some of the issues raised by the Web, it is important to also have a closer look at some specific online applications and mechanisms, as the Web is indeed a *web* of applications and networks. This will be done in the following chapters.

Chapter 5

Social media

Contents

5.1 Introduction

Everything that human beings are doing to make it easier to operate computer networks is at the same time, but for different reasons, making it easier for computer networks to operate human beings.

George Dyson, *Darwin among the Machines*, 1997

Around the turn of the millennium and onwards, we can see a change in the Web's character that is still ongoing today: the bidirectionality of the Web became a more prominent feature in online applications (Raffl *et al.*, 2011, p. 608). At least partially driven by the view that "changing the world for the better and making money aren't mutually exclusive" (Schäfer, 2011, p. 31), the Web started to gain momentum as a medium for social interaction and collaboration. Often, this is referred to as a shift from 'Web 1.0', in which the Web was a relatively static consumption environment, to 'Web 2.0', in which the main use of the Web is interaction and communication (see e.g., Cormode & Krishnamurthy, 2008; Beer, 2009; Raffl *et al.*, 2011). This resulted in a change in web pages, as they went from presenting predominantly text-based and relatively static Web content (discussed in chapter 4), to generally more pictorial and dynamic content that allows "fluid interactivity such as rearranging and editing data, manipulating graphics, or playing games" (Jamieson, 2016). Meanwhile, the ongoing implementation of WYSIWYG-publication applications continued to lower the barriers to online publishing. Online publishing became a matter of typing a sentence or uploading an image, and clicking 'OK'. The combination of the publishing affordances and the increased emphasis on user communication and interaction turned *user-generated content* into a prominent focus of Web use (Beer, 2009, p. 986). A key role in this character change of the Web is played by social media sites.

Social media sites are, as their name already suggests, online applications that offer social interaction between users. They are mediating infrastructures that shape the social acts performed on them (van Dijck, 2013, p. 29). Social media usually offer some form of connecting a user to other users and allow them to bidirectionally communicate. Van Dijck points out several types of social media: (1) social network sites (SNS) like Facebook and LinkedIn — applications that promote social or professional interpersonal contact; (2) sites focused on user generated content like Youtube and Instagram (although Instagram seems to be turning more and more into a type 1 social media), which promote creativity and the exchange of amateur and professional content; (3) Trading and market platforms, like Etsy.com; and (4) websites for playing games (van Dijck, 2013, p. 8). The most prominent social media, and the most relevant for this study, are SNS. These can be defined as "Web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse

their list of connections and those made by others within the system" (Ellison & boyd, 2007, p. 211).

By mediating social interaction, social media are bound to entail the encoding and transmission of personal information. It is therefore important to explore how social media affect the online informational persona and what complications this may entail. I will explore this in this chapter.[1] As I did in the previous chapter, I will trace the impact of the technological mediation on the online assimilation of personal information in three directions that relate to main elements that shape the perception of the informational persona (see section 2.3): the production of information (and thus the content of the informational persona), the presence of the informational persona, and the composition of its publics. Lastly, I will conclude this chapter by reviewing how the assimilation of personal signifying objects by social media can complicate the portrayal of an individual by her informational persona.

However, the analysis of the impact of the mediation by social media is not a straightforward task. There is a high diversity in social media and their target groups, intended functions, and goals. As a result of these differences, also their design, mechanisms, and use vary. Moreover, many social media change their architecture over time and continue to do so in rapid succession. Because examining all the social media and their differences in detail is a research topic on its own, I have decided to focus on the architecture of one of the most popular services: Facebook[2]. Although Facebook is not exemplary for all social media, it does portray several main features that can be found in many social media. Moreover, because I cannot keep up with all the latest updates, I ask some lenience of my readers with regard to settings and practices that may be outdated by the time this dissertation is printed, and to keep in mind the general lines of thought that I present here. With this main example in the background (and sometimes others where it is called for), I discuss the implications of the mediation by social media for the online informational persona.

5.2 A social media example: Facebook

In this section, I will briefly describe the general use and features of Facebook.[3] This section is primarily meant for readers who never ventured unto Facebook or a similar social media application. Readers familiar with the use of social media and concepts like 'feeds', 'walls', 'tags', and 'likes' can skip this section.

Facebook is a social media application that offers users an online space where they can interact and share information with other users. It is used for a wide range of social interactions, varying from professional and commercial to intimate

[1] Other topics with regard to social media sites, like the use of social media for surveillance by governments, or the consequences of the aggregation of data for commercial purposes, lie outside the scope of this research.

[2] Facebook had roughly 1.52 billion daily active users around December 2018. See `https://newsroom.fb.com/company-info/`, last accessed 16-02-2019.

[3] This analysis of Facebook is based on Facebook as it was around 2018-2019.

and household interactions. In Europe, Facebook is widely used. At the end of 2018, roughly half of the European population had a Facebook account.[4]

The Facebook application is made accessible for various types of devices and is free of charge for users. Despite this free of charge use, Facebook is run by a for-profit corporation. Its main revenue derives from selling advertisement placement.[5]

The idea underlying Facebook is that people should be 'rewired' to share and consume more information and provide for an accurate and transparent self-representation (Mitchell, 2014). This philosophy is tightly embedded in Facebook's setup and policies: users need to create a profile account with, at the minimum, a valid email address. For the user profile Facebook enforces a 'real name policy', requiring users to use their 'real ID' name.[6] Facebook is open for everyone above the age of 13 (with potential deviations based on local laws).

Users interact with Facebook through this profile account. With her profile, the user can connect to others or interest groups and keep a tab on the latest developments. Also, it allows the user to set different access and privacy options so that the user can share information with specific groups or connections.

It is possible to access some of the content on Facebook without an account, but the access is limited and forged into a cumbersome experience where the user is systematically pressed to create an account by a huge white banner in the middle of the screen asking the user to log in or create a Facebook account.

5.2.1 Main features

Facebook has various features that shape the information flow between the user, others and the social medium. To give the reader a general impression, I will briefly describe the main features that are relevant in the light of this study. Next to these, Facebook also offers services like instant private messaging, but I will not discuss these services here.

Walls Facebook is built up from 'walls', which are personal 'timeline' pages of users or interest groups. On these walls the users themselves or others can upload text, images, videos or sound files. The amount of content that users can upload on their Facebook page is almost unlimited.

News Feed The homepage of a user displays a 'feed' to her. The 'feed' is a list of signifying objects from different locations aggregated in one point. Here, the users can quickly see the latest or most popular posts by their connections without having to browse to content. As such, the feed allows the user to passively receive

[4] `https://www.statista.com/chart/16256/facebook-users-in-europe/`, last accessed 15-04-2019.

[5] Facebook, *Facebook Financial Report 2018*, available at `http://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2018.pdf`, last accessed 19-4-2019.

[6] Facebook, *What names are allowed on Facebook?*, `https://www.facebook.com/help/112146705538576`, last accessed 19-4-2019.

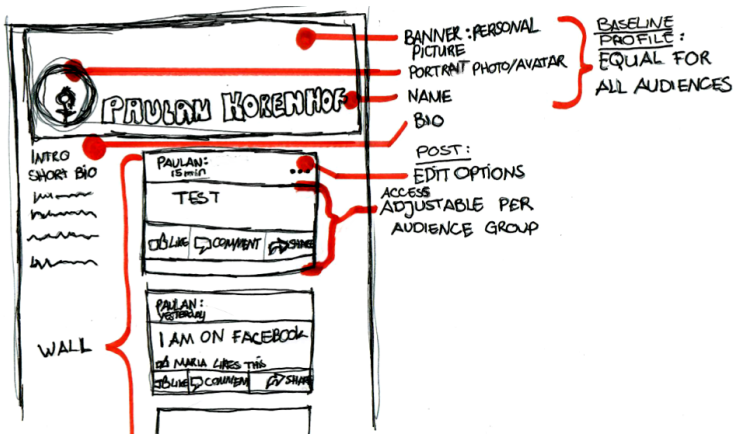


Figure 5.1: Anatomy of a Facebook profile page

signifying objects; the information is being 'fed' to her by means of algorithms. I will discuss the mechanisms of the feed in section 5.5.2.

Like buttons, emoticon buttons, and comments On Facebook, users can react to published signifying objects by (1) clicking a 'Like' button as sign of approval or enjoyment, and/or express their feelings by clicking and selecting an emoticon, or (2) comment on the content by typing in text, or uploading a picture or video file.

Tag The tag-function allows users to 'tag' other users or themselves in photos, thereby linking a specific user to a certain part of a photo (usually a face).

5.3 Mediating platforms

In this section, I take a closer look at what kind of applications social media sites are, how they are controlled, and what implications this has for the users they mediate.[7] This will provide a general background for the upcoming sections in which I examine the content of the informational persona on social media, its presence, and the composition of its audiences.

With a website as a key element of their architecture, social media sites share many affordances with 'basic' websites as discussed in chapter 4; they can almost

[7]In this section I highlight what I take to be the main differences between social media and 'basic' websites taken as relatively static (HTML)documents (see chapter 4). However, it is important to note that 'basic' websites, as well as social media, come in all sorts of variations that can share more or less similarities depending on a particular case. Tracing the all the possible (technical) differences between social media and basic websites is a research on its own and lies outside the scope of this study. For practical purposes, I therefore present a somewhat simplified and streamlined view.

instantly be accessed from anywhere, while the content can easily be copied and spread. However, social media also give rise to novel dynamics in the online information flows. The core of this difference lies in the control over the website. Websites, as discussed in chapter 4, are in general focused on sending a message from the website controller to a visiting user, and thereby establish a line of action in one direction between two parties:

$$\text{website controller (sender)} \rightarrow \text{users (receivers)}$$

Social media diverge from this model by separating the role of the sender from that of the website controller. They step into the middle by offering users the tools to publish, but not the control over these tools. Moreover, they are focused on realising two-way communication between users. With this, social media take on the role of mediating website controller in the informational interaction between users:

$$\text{user (sender/receiver)} \leftrightarrow \text{website controller (social media)} \leftrightarrow \text{users (receivers/senders)}$$

The typical characteristic of social media is thus that they mediate informational interactions between users. Their content is generated by users; without users and user generated content, social media are empty shells. The space in which the user-actions take place, as well as all features with regard to the encoding, storage and retrieval of content on social media sites are determined by the controller of the social medium (hereafter: medium controller).

On the level of a mediating controller, there is a sliding scale between 'basic' websites that are constructed by means of push-button publishing and social media applications. In some cases of 'regular' push-button publication web pages, the users may be offered so little control over the tools to publish, that the web page may seem to be more resembling the control structure of social media, than that of a regular web page. Broadly speaking, the difference between the two, is that the role of push-button-publication controllers tend to be rather passive (of course, this does not say anything about the role of the mediating technology), while social medium controllers tend to actively do things to the content that users publish: they process it in feeds, add content, delete content, etc. However, the reader should keep in mind that this is, indeed, a sliding scale and that my description is not meant as a strict demarcation between social media and other types of online applications.

Due to their communication-mediating role, social media tend to have highly flexible content. The driving force of this flexibility is the incorporation of 'AJAX' (Jamieson, 2016). AJAX stands for Asynchronous JavaScript And XML. AJAX is a method that "allows Web pages to be updated asynchronously by exchanging data with a Web server behind the scenes. This means that it is possible to update parts of a Web page, without reloading the whole page"[8]. The implementation

[8]W3Cschools, `http://www.w3schools.com/xml/ajax_intro.asp`, last accessed 02-01-2017.

of AJAX eliminates the need for users to refresh the page in order to gain access to the newest content on the site, thereby smoothing the process of information-interaction. The flexibility of content and use of social media is generally further extended by built in application program interfaces (APIs) that run on top of the social media sites (van Dijck, 2013, p. 8). These APIs allow for an easy development and implementation of applications like games, chat and email in the main social media site. Social media sites are thus often multifunctional and function as a 'platform' where myriad services and content can be offered (Gillespie, 2010, p. 348).[9]

Because the social media site *is* the point of contact between users, the user interaction will be heavily affected by the design of the medium's interface. As the medium controllers control the architecture and build their ideas and values into the design, their code regulates what users can and cannot do on the medium. The consequence is that the user interactions take place in an architecture that is shaped by the ideas and values of the medium controller. To give an example, a user cannot sign up with Facebook without providing a first name, a surname, mobile number or email address, a date of birth and a gender. Hereby Facebook establishes already a significant base identity. Moreover, by requiring a mobile phone or email, Facebook steers users towards having only one profile per person. Another example is the incapability of users to prevent their updates shown on their timeline from showing up in the feed (I will discuss the feed in detail later). Any interaction between two users on a timeline is thereby in principle automatically broadcasted towards a larger audience via the feed. This reflects Facebook's aim to engage as many users as possible.

By building rules for user engagement in the technological architecture, the medium controllers apply a form of *techno-regulation* (cf. van den Berg & Leenes, 2013). The autonomy of the user is restricted to what is offered by the medium controller. If the user disagrees with a particular restriction or platform mechanism, she remains with the choice to either accept it or to refrain from using the medium.[10] As such, the medium controller has a dominant and active position in the shaping of the information flows on social media. It is therefore important to take a closer look at the manner in which medium controllers tend to fill in this role.

Currently, many social media are under corporate control. The result is that business models now often underlie the construction choices on the social media's infrastructures. Offering users their services for free, these controllers make profit by exploiting the users' cognitive time and energy, turning their cognitive capacity into an important productive resource (Lazzarato, 2014; Berardi, 2009a; Virno,

[9]Gillespie points out that the term 'platform' can have various (politically coloured) connotations (Gillespie, 2010). I would like to refer readers who wish to know more about this, to Gillespie's article *The politics of "Platforms"* (Gillespie, 2010).

[10]This is not a two-way street: a user who chooses not to use a certain social medium can still be used by that social media. For instance, many websites place Facebook cookies on users' devices irrespective of whether the user has a Facebook account (Roosendaal, 2010). However, because this has little impact on the relations between users, I have decided to leave this specific discussion out of the scope of this study.

2003; Moulier-Boutang, 2011; Berardi, 2011; Stiegler, 2010a). The information that users share is often used for advertising goals (Cohen, 2013). By spending attention, interacting and generating data the user 'pays' for the service (Terranova, 2012). On social media, we thus see the emergence of a particular kind of user, namely one that is simultaneously "a resource provider, a product, and a consumer" (van Dijck, 2013, p. 170). Users provide resources by encoding signifying objects, and consume the objects published by others. All the while, the user generates data and spends attention, which are used by the medium controller as a product to sell for profit to third parties.

Given this revenue model, corporate-run social media have an interest in attracting as many users as possible, while prolonging the time they spend on the platform and increasing their online productivity and activity by communication, content production, creative actions, and the establishing of connections and communities (Fuchs, 2013, p. 105). The size of the user base and the level of their engagement is critical for the success of corporate run social media that depend on selling user attention. Facebook emphasises: "*If we fail to retain existing users or add new users, or if our users decrease their level of engagement with our products, our revenue, financial results, and business may be significantly harmed*" [emphasis in original text].[11] Because the size of the user base matters, social media have an interest in keeping accounts in their system, even if users do not engage with the medium anymore (Leenes, 2009, p. 50). The result is that media controllers are inclined to steer users towards the retention of their account in some form, like offering options for temporary deactivation instead of irreversible termination.

In order to attract and maintain as many users as possible and maximise engagement and information sharing, social medium controllers tend to employ techniques like *nudging*, *persuasion*, and *gamification*. These techniques are intentionally built into the architecture of the mediating technology in order to influence users to behave in a particular manner (cf. Thaler & Sunstein, 2009; van den Berg & Leenes, 2013). I will briefly discuss these three techniques.

While *persuasion* may initially seem like an action between human agents, several authors have described how it can be embedded in the technological architecture that establishes human-computer interaction (see e.g., Fogg, 1999; Harjumaa & Oinas-Kukkonen, 2007). It is a form of attempted influence to steer or alter user behaviour and attitude (Oinas-Kukkonen & Harjumaa, 2008). Persuasion in a technological form entails the presentation to users of a relatively explicit and clear choice with an outcome favoured by the designers (van den Berg & Leenes, 2013). An example of this is a banner that overlays the content of a website and requests the user to log in. The user cannot ignore this banner; she either needs to log in or keep clicking the banner away. The technology thus tries to persuade the user to log in or create an account, while it does not take away her choice not to do this.

Nudging works more on the unconscious level: it works by shaping the context of people's choices in such a manner that it "alters people's behaviour in a

[11] Facebook, *Facebook Financial Report 2018*, available at `http://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2018.pdf`, last accessed 19-4-2019.

predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2009, p. 6). While nudge requests may be prominently placed in the interface, they do not hamper the user in her actions and choices. However, they do affect user behaviour. The concept of nudging relates to the non-neutrality of technology: as Weinmann, Schneider, and vom Brocke argue, "there is no neutral way to present choices" (Weinmann *et al.*, 2016, p. 433). Incorporating nudging techniques in online applications entails the shaping of the environment in such a manner that users' choices are unconsciously affected towards realising behaviour favoured by the designer. Nudging offers the user a limited choice, but does so less explicitly than persuasion techniques (van den Berg & Leenes, 2013). Examples of online nudge techniques are the display of a password strength in order to nudge users into using stronger passwords and the use of default settings to nudge users into an opt-in instead of an opt-out (or vice versa) (Weinmann *et al.*, 2016, p. 433).

The third technique, *Gamification*, is "to use elements of game design in non-game contexts, products, and services to motivate desired behaviour" (Deterding, 2012, p. 14). Gamification focuses on stimulating user engagement and works with activity loops that revolve around action, feedback, and emotion (Ibanez *et al.*, 2014, p. 291-292). By rewarding users with points or the like, the technology can play in on emotions and motivate users to 'play the game' and keep performing certain actions for more points. The technique is for example user to motivate learning behaviour in students (see e.g., Ibanez *et al.*, 2014)

The offering and presentation of choices in an online architecture (or any for that matter) is thus inherently non-neutral. By offering users a set of choices and implementing techniques like the three above, the media controller can highly affect user behaviour. For example, the information sharing behaviour of individuals can be affected by shaping the technology in such a way that it suppresses privacy concerns (Acquisti *et al.*, 2015, p. 509). The various techniques give rise to a hybrid intentionality with different user and technological intentionality-ratios; in some cases the user is given almost no choice (for example, she cannot create an account without giving an email address), while in others she is given a limited choice, or an extensive choice. The more a user can freely choose and act, the more strongly she can express her own intentions in this hybrid intentionality. However, even choice settings that allow much room for the expression of user intentionality, can have a deep (unconscious) effect on the choosing user by being presented in a particular manner. The most striking example of this, is the default setting. By installing certain features as default settings, the website sets a standard for interaction on the platform, and burdens individuals with divergent preferences to change these settings (see e.g., Gross & Acquisti, 2005; van den Berg & Leenes, 2013; Acquisti *et al.*, 2015). Users are often inclined to accept the default setting, because it "is convenient, and people often interpret default settings as implicit recommendations" (Acquisti *et al.*, 2015, p. 512). With the default settings, the controller can exploit users' cognitive biases and nudge them into performing particular behaviour (Weinmann *et al.*, 2016, p. 434). Default settings thus have a strong influence on user interactions and in turn affect the user's norms. Next to

employing nudging techniques like default settings, the social media architecture often also has built in certain persuasion techniques, and gamification elements.

With the use of behaviour influencing and regulating techniques, the medium controller builds her own norms into the social media architecture. As these 'coded' norms are pressed onto users, social media brought about changes in Web culture by giving rise to a new standard of what is considered 'normal' (van Dijck, 2013; Wittkower, 2014). One of the biggest norm shifts brought about by social media is the shift from the use of the Web for relatively anonymous communication and interactions to patterns of communication where "individuals are increasingly known, and in fact willingly share a lot of their personal information online" (Sparrow *et al.*, 2005, p. 283). As such, social media led to a normalisation of being online identifiable as a particular offline individual.[12] The use of real names on Facebook is an example of this. The result of this norm shift, is that much of the online information clearly identifies a particular offline individual as its encoder — thereby tightening the informational relation between the online and the offline individual.

Lastly, it is important to touch upon the individuating character of social media architecture. In general on social media, every user is represented by a profile, and can only add personal information and perform actions within certain technologically predefined bounds rooted in this profile. An example of this is that all Facebook profile pages have the same layout template. The user can conveniently fill in the information sections, upload images to fill up the banner-block and the avatar, and move in predefined action-paths. By letting users act only within clear bounds and a predefined uniform action template, the user's informational persona can be compartmentalised efficiently and translated into coded material that is easy to process and control by the medium controller (Bucher, 2012, p. 1171). By homgenising and classifying users, the technological architecture strengthens the control of the medium controller and accommodates the production of value (Negri, 2005; Terranova, 2012, p. 107). In this process, the individual is divided and materialised into chunks of code,[13] as the masses are turned into data banks (Deleuze, 1992, p. 5).

All these processes take place in an environment that is relatively opaque to the user; while the user is made highly visible, the functioning of the architecture and the medium controller's role herein, are generally hidden from the users by the interface — much like a one-way window. This serves the purposes of the media controller as users "are not supposed to understand that we are the product of marketers as much as we are the market" (Vaidhyanathan, 2008). The medium controller wants the users to relax, and feel comfortable to reveal information about themselves (Vaidhyanathan, 2008). On social media, the extent of the power of the media controller is thus hidden by an asymmetry between the visibility of the media controller and the user (Trottier, 2011).

[12]The changes brought about by social media platforms are not welcomed by all users and some agents decide to 'push back' against the constant expectation of pervasive connectivity and availability (Morrison & Gomez, 2014).

[13]Referred to by Deleuze as 'dividual' (Deleuze, 1992, p. 5).

5.4 The production of social media content

With the characteristics of social media sites as discussed in the previous section in the background, it is now time to look at the production of personal information on social media, the presence of this information, and the composition of its publics. In this section, I will start by examining the production of personal content.

On social media, the how, the who, and the what of the content production are closely intertwined. However, in order to give some structure to this section, I will maintain the elements as used in the previous chapter, but in a different order. Due to the typical character of social media, I will start with the 'who' instead of the 'how'. Next, I will examine the 'how' of the production of personal information and the *what* that is produced. Lastly, I will summarise the main points and their impact.

5.4.1 Who: choices of publishers

The 'who' that can encode content on social media is determined by the medium controller. Compared to regular web pages, we see on social media generally a restriction of the 'who' that can publish because the publishing of content is often restricted to users with an account. This account comes with a profile, that, at the minimum, consists of an identifying element like a name and a profile picture or avatar. It plays a crucial role in interactions on social media, because it is for users generally not possible to interact outside of this profile (only watching is sometimes allowed without a profile, but under certain conditions). With this, the users are individuated. Many social media aim to shape the profile in such a manner that it establishes a unitary profile per user, by for example allowing only one profile per user account and email address (Wittkower, 2014).

Next to the users, also the medium controller can publish content on the platform, or set up the platform to automatically publish certain content. Since the platform is under the control of the media controller, the media controller has a high discretion in her publishing choices. Additionally, some third parties may be able to publish on the platform, if they have an agreement with the medium controller to do so.

5.4.2 How: means of production

The means of publication on social media are in the hands of the medium controller. Because social media rely on their users for content and revenue, its architecture tends to be designed to simplify, and even promote, social expressiveness (Berardi, 2009b, p. 153). This is done by the embedding of persuasion, nudging, and gamification techniques in the architecture, and by making the production of content as easy as possible and open for a wide range of users. In order to achieve the latter, social media are generally equipped with a WYSIWYG-format interface to allow for quick and simple encoding that requires little know-how (see section 4.3). This interface is often set up for the publishing of miscellaneous types of

objects, like images, text, and video- and sound files. However, the implementation of a WYSIWYG interface on social media generally has a typical character: it requires the user to fit the content that she wants to publish into a particular preformatted layout that forces users to encode rather uniform objects (van Dijck, 2013, p. 161). For example, Twitter has a built in limit to 280 characters per 'tweet' in a typical layout. Anything that exceeds the 280-character limit needs to be split over separate entries. The social media architecture thus gives rise to an industrialised information publishing platform that produces a particular style of signifying objects.

Moreover, because the signifying objects on social media are meant to be the focal point of social interaction, they are generally equipped with a certain 'annotability'; the social media architecture allows others, as well as the user herself, to add comments or other signs to these objects (Lemmens, 2014, p. 7). By allowing, and even inviting, the audience to publish their reactions on the content, social media transform the traditional division between producers and consumers. As multiple users encode reactions to each other, we see the creation of a multicomposed signifying object in which the various reactions of users are entangled in one evolving signifying object. In being annotated, the object becomes a 'live' negotiation of its meaning and value in an interplay between users and mediating technology. The meaning of the signifying object is produced at the crossroads of "information processing, software dynamics, linguistic articulation, and cultural practices" (Langlois, 2013, p. 91).

Additionally, the policies of the media controller affect the content that is encoded. These are policies that determine what kind of content is prohibited and what is allowed, but include policies that see to the establishment of the user profile. An example of the latter is Facebook's real name policy. With such policies a medium proscribes a certain standard about what a user's 'authentic' identity can be (Haimson & Hoffmann, 2016). These policies tend to be backed up by a certain level of enforcement. For instance, an accusation of not abiding by Facebook's real name policy can lead to a suspension or ban of the user's account. Such policies highly affects the materialisation of the user into a particular identity representation on the platform. As these policies highly affect what identity on the medium *can* and *should* look like, they can complicate the use of the medium for users with a non-normative or fluid identity (Haimson & Hoffmann, 2016).

Lastly, it is important to briefly touch upon the devices that users can use to encode content on social media. Publishing content on social media is generally available for a variety of devices (personal computer, smartphone, tablet). However, especially the manner in which social media are integrated in smartphones is important to point out here. Smartphones generally provide applications that allow the rapid publishing of photos, videos and sound files on a social media platform: a photo that was just taken, can be published in mere seconds. This speed and simplicity of uploading and publishing mechanisms can easily trigger spur-of-the-moment actions with little thought or reflection (Wang *et al.*, 2011). What adds to the ease of publishing personal information, is that despite the fact that on social media users tend to have a certain visibility due to

their profiles, their interactions are still interfaced and they remain susceptible to some degree of the disinhibition effect (see section 4.3.3).

5.4.3 What: the personal product

The 'what' that is encoded on social media takes shape at the crossroads of the 'who' and the 'how'. Because content on social media can cover any topic, I will focus on the general tendencies with regard to the content that is encoded, and in particular in relation to personal information. Building further on the impact of the 'who' and the 'how', I will draw a general picture of the 'what' that is encoded.

First off, the content that a user produces on social media, depends of course highly on the user herself. On social media, researchers found for example differences in social media posting behaviour between users of different age groups (Pfeil *et al.*, 2009), of different social and cultural backgrounds (Kim *et al.*, 2011), and between users with different personality traits (Lee *et al.*, 2014). However, the user behaviour on social media always takes form under the influence of a hybrid intentionality. The social media architecture, its general use, as well as the policies of the medium controller, also affect the content users encode. Various social media are therefore likely to affect the content in different directions. For instance, users tend to disclose more intimate information on Facebook than they do on Twitter (Choi & Bazarova, 2015). Despite these differences, I will discuss the general characteristics that are relevant for the content on many social media.

The most crucial element of social media that affects the 'what', is the profile. The profile entails a representation of the user *through* which she interacts. It individuates the user and is expressed by an identifying element (user name, profile picture) that is automatically attached to all the signifying objects that the user encodes on the platform. The consequence is that all these signifying objects refer directly back to the user who encoded them. The user profile itself is often in an extended form presented to users by means of a single profile web page. With its prominent role and lay-out, the profile page constructs a representation of the user's identity by a set of signifying objects produced by the users, others, and the platform. The profile page functions as proxy for a specific and often identifiable offline person, and imbues her with a continuous personal presence on the platform that places user always within reach of her connections. As such, the profile establish a certain degree of what DeVito, Birnholtz, and Hancock call 'identity persistence', which is "*the extent to which a platform affords the identification of content with an individual persona over time*" [emphasis original] (DeVito *et al.*, 2017, p. 742). The creation of a profile is therefore an important part of participating in social media: it "is an explicit act of writing oneself into being in a digital environment" (boyd, 2010, p. 43). The construction of a profile can be a meaningful way for individuals to play around with distinct roles and develop their self-identity (cf. Wandel & Beavers, 2010) — though this necessarily takes place within the boundaries defined by the medium controller.

The profile page has a feature that allows the user, as well as others, to publish information on this page by means of push-button-publishing. In the case of

Facebook, the feature that enables this, is 'the wall'. Even if this feature is used to directly contact another user, the wall "— as its name already suggests —, has a social function that extends beyond the two primary actors in the communication" (Leenes, 2009, p. 54). The wall displays its content to the audiences of the user. As such, the wall encourages openness and expresses social relations and the user's place in the social network to a wider audience (Leenes, 2009, p. 54). Due to the visibility of these publications, the use of social media for information sharing tends to have a broadcasting character (Berger, 2013, p. 294). This affects the act of encoding and the content that is encoded. Because the user is publishing information to a potential broad and invisible audience (I will get back to this in section 5.6), she is unable to focus her choice of content on a specific other, unless she addresses this other directly. With only little idea of to whom she is talking and therefore which topics to address, the user is more likely to focus the content she uploads on presenting her self (Barasch & Berger, 2014, p. 17). Additionally, users are more inclined towards casual communication in this setting compared to directed communication (Condella, 2010, p. 116). On social media, we therefore see the encoding of a stream of personal signifying objects referring to what users did, what they read, what they liked, how they look, etc.

Moreover, the social broadcasting character establishes a low threshold for social interaction: it allows users to ventilate emotions and the like, and invites social support, but without giving a user the feeling that she burdens a specific individual (Berger, 2013, p. 294). The posts on social media give others "a window into my own world—be it what I'm doing, how I'm feeling, or what I'm thinking—in a way that does not intrude on the time or space of others, but allows them to discover these things for themselves and at their own leisure" (Condella, 2010, p. 116). Combined with the affordances of ubiquitous and spur-of-the-moment publishing options, we see on social media therefore a focus on a 'what' that is generally personal, and may also be emotional or sensitive in nature. Users even report to have shared sensitive and/or strongly sentimental content — content they later regretted —, while being heavily emotional, intoxicated, or when they misjudged the context in which they were posting (Wang *et al.*, 2011). However, it is important to note that in general, social media users encode content that entails a positive self-presentation, and some even present an ideal self (Seidman, 2013; Lee-Won *et al.*, 2014). Users feel that negative content may entail risks for their self-presentation, while positive content is generally taken to be constructive for their social image (Ma *et al.*, 2016).

While the content on social media has had all sorts of forms and styles, there is one particular type of content that I find worthwhile to specifically touch upon due to the content's highly self-referencing character (contrary to for example food pictures): 'the selfie'. A selfie is a self-portrait made by the referent with the camera of a mobile device at an arms length. It is a visual self-reference that expresses a self-presentation: the user intentionally (although not necessarily well thought through[14]) takes a picture of herself to present herself in a certain manner

[14]See e.g., Bryn Lovitt, "Death by Selfie: 11 Disturbing Stories of Social Media Pics Gone

or in a particular context.[15] With the selfie, the user wants to associate her own image with this particular context (Leone, 2018). Users post selfies in order to seek attention, communicate to others, for entertainment, and/or for archiving purposes (Sung *et al.*, 2016). Users even state that they post selfies in order to be acknowledged or have their existence reaffirmed by others (Sung *et al.*, 2016). They thereby consciously employ selfies to engage in reflexive identity construction with others. Moreover, the selfie itself can work as a kind of reflexive identity mirror for the referent: as a materialised representation of the self seen through the lens of the mediating technology, the selfie works as a self-affirmation for the referent by showing to herself that she is indeed a particular kind of person (cf. Toma & Hancock, 2013; Leone, 2018). Much more can be said about the selfie. However, I will need to leave it at this because the selfie is specifically tied to the use of particular input devices and is a research topic on its own.

Despite the fact that users on social media often focus on themselves when publishing content, they can also intentionally or accidentally create references to others. As others cross paths with the user, they can become (co-)actors in the publications of the user: whether it be intentional in a joined event (e.g., a post about going to dinner or a movie together), as a random stranger accidentally captured in the background of a photo, or as a stranger intentionally recorded because he raised the interest of the user. The social media architecture normalises the inclusion of others in content by asking users to include and identify other people in their posts by means of 'tagging' these others. The 'tagging' establishes a link from the content to the profile of the tagged user. As such, on social media we see "the near-universal practice of posting content about others on one's Web page" (Nissenbaum, 2010, p. 60). The effect of this practice is that a part, or even a significant amount, of the information relating to a particular referent may not be intentionally shared by the referent herself. The referent may not even be aware of the creation of a signifying object referring to her.

Adding to the user generated content, we can also see the addition of new kinds of content by the social media application itself. One important example is the display of the user's connections. These connections are often automatically publicly displayed on the user profile, such as a line under the user name stating "265 friends". The display of a user's connectivity can function as her social capital: it can signal popularity and social importance. I will discuss the mechanisms and implications of this in section 5.6. Another relevant example is that a platform can have automatic publishing mechanisms in place that encode and publish certain information about the user's activities on the platform without the user's interference (and often also consent). Usually these are short signifying objects like "A is now connected to B", "A is interested in this event", and "Today is A's birthday!". These publications can in turn spark new interactions, and thus the generation of more content. The choice of content for these automatic

Wrong", *Rolling Stone*, 2016. `https://www.rollingstone.com/culture/culture-lists/death-by-selfie-11-disturbing-stories-of-social-media-pics-gone-wrong-15091/misstep-at-the-taj-mahal-28874/`, last accessed 13-04-2019.

[15] Some users share as many as 650 selfies a month (Sorokowska *et al.*, 2016).

publications is fully dependent on what is built into the platform architecture. As such, these publications reflect a strong expression of technological intentionality. Whether and to what extent users can exercise control over these automatic publications depends on the options offered by the platform.

Lastly, companies and organisations can seek to make use of user generated content on social media for advertisement purposes. They can invite users to publish their interest in a product or in the company itself. Companies and organisations also can have social media functions built into their web pages, allowing users to easily share, 'like', or 'tweet' their interest in the company or product with a single action — often in turn for a chance to win some prize. Such a built in feature "facilitates and normalizes linking corporate-owned sites about one's interests to one's own account or profile" (Stanfill, 2015, p. 1066). With help of the user, the user profile page can thus be used as a signboard for advertisement. However, it can also be the case that a company actively scrapes user information and incorporates this into an ad that is presented to users.

5.4.4 The who, the how and the what of social media content

The production of information on social media results from various mixes of a hybrid intentionality between users, audiences, third parties and the mediating technology. This gives rise to a wide array of objects in which different agents predominate. What these objects have in common, is that they have a personal character.

Due to the social and highly interactive character of social media, the user is at the same time publisher, referent and observer. At the minimal level, she is the co-referent of every signifying object that she publishes, because her identifying profile elements accompany her every comment and post, even if she publishes about others on their profile pages.

While users play the leading role in the creation of content on social media by creating content about themselves and often also about others, they do so in a hybrid intentionality with the platform architecture. The platform architecture and policies guide the creation of content into certain formations and along certain action paths. Users need to translate their identity into prefabricated formats and sections that they can fill out, while complying with the policies of the medium controller. The medium controller and architecture thus both express a significant intentionality with regard to the 'what' that their users encode.

5.5 The presence of information on the platform

As applications on the Web, social media in principle give rise to some of the same affordances with regard to the presence of personal information compared to regular web pages. However, with their particular characteristics as discussed in section 5.3, they also give rise to some different information flows. In this section,

I will examine how social media affect the presence of information for users on the Web, within social media platforms themselves, as well as within the platform over time.

5.5.1 Presence within the Web

Social media like Facebook take in a particular position on the Web: they are platforms that offer a central top-down organised access point for engaging with the Web through a cluster of connected services (e.g., chat, mail, information access, games). This centralisation of services in one website, allows the platform to function as a gateway and identity provider for users (van Dijck, 2013, p. 64). The user can surf and access content through the social media homepage. By offering this cluster of services in one access point and connected to one user account, social media bring about a shift in the Web's landscape: they form silos on the Web that is relatively isolated from other content. The user is offered an often personalised experience of the Web in the platform's own relatively autonomous information ecology that tends to lock users "into closed, centralised walled gardens" (Lovink, 2016, p. 37-38). This lock in is even stronger on smartphones than on desktops or laptops: on the smartphone the interaction is fully submerged in the application controlled by the medium controller, instead of first through a Web browser. However, it is important to note that users are likely to use multiple social media applications. For example, users may use Facebook for their social connections, LinkedIn for their professional contacts, and Twitter to be involved in public discussions. As users switch between applications, they switch between silos.

Due to this lock in and the top-down centralised infrastructure of the social medium, the medium controller has a strong influence over the presence of personal information on social media and their audiences. While many users are aware of the risks that come with this and have privacy concerns, they still choose to engage in social media (see e.g., Barnes, 2006; Taddicken, 2014; Acquisti *et al.*, 2015). Peer pressure and social motivation prompt people to use these applications and maintain a presence on the platform (Leenes, 2009; Brandtzæg & Heim, 2009; Berardi, 2009a; Quan-Haase & Young, 2010; Wittkower, 2014).[16] Users who lack access and visibility on particular platforms run the risk of lagging behind or even fully missing out on social and professional events. The popular use of a social media giant like Facebook, contributes to the pressure on individuals to have a presence on the platform. The more users, the higher the utility of the platform, and the more difficult it becomes to leave. User choice with regard to the medium thus often boils down to an opt-in/opt-out choice between compliance with the platform and creating a profile, or being excluded from certain content, groups and connections of interest. Access in this manner is heavily intertwined with social media's regime of user individuation and control: "a regime in which the user is habituated, on the pain of exclusion from social worlds, to surrendering

[16]This in itself comes with a risk for users: the ongoing interaction with others established by social media can contributes to the likelihood of its users getting a burnout (Zivnuska *et al.*, 2019).

the elements of their personality — identity, creativity, sociality — to enhance the circulation of capital" (Dyer-Witheford, 2015, p. 93). Social media therefore give rise to what Deleuze referred to as *societies of control* (Deleuze, 1992). Users on social media are being monitored and steered in the information flow by a systematic calculation of their preferences in a feedback loop, while codes "mark access to information, or reject it" (Deleuze, 1992, p. 5). The medium controller thus exercises a certain power over the users through the medium's architecture: users can only use the social media as the architecture allows. Fuchs therefore questions — or rather criticises — the notion 'participatory' that is often used to describe the role of users in social media (Fuchs, 2013, p. 98). Users may attempt to participate with each other *on* the social media platform, but they do not participate *in* the social medium. The technologically mediated sociality of social media thus brings a certain tension with it: on the one hand individuals want to participate in the online social interaction, while on the other hand this participation means that they will confide personal information to a mediating technology that is controlled by a third party that is generally motivated by its own interests. For control over their information and its visibility on the platform, users are dependent on the options offered by the social medium. In the next section, I discuss how this visibility and attention is steered by the platform.

5.5.2 Presence within social media

While the content of signifying objects on social media depends for the majority on choices of users, their presence is highly shaped by the social media architecture. The presence of information over time and space depends on the access options, layout, and infrastructure features of the social media. The medium's architecture shapes the information flows and influences how its users experience the meaning of the content by setting up the parameters of the communication and information exchange (Langlois, 2013, p. 98). The control over the medium thus entails a "management of flows of meaning" (Langlois, 2013, p. 91). The two main levels on which the architecture affects the presence of information within the platform for a particular user, are the medium's connective mechanisms, which I will discuss in section 5.6, and the information flow given a particular set of user connections, which I will discuss in this section.

The main feature employed to manage the 'flow of meaning', is the 'feed'. A feed, like Facebooks 'News Feed', helps users to cope with the often massive offer of information by streamlining it. The feed is commonly displayed on the first page that users will see when they log in. Because of its visibility and role on the platform the feed is one of the primary features of social media like Facebook (Treadaway & Smith, 2012; Bucher, 2012). In a feed, new and updated content is aggregated from user and group pages by automated processes and pushed towards the user on a single page by 'feeding' her a stream of signifying objects. Feeds unburden users by saving them the time and effort of having to manually visit all the profile pages of their connections or interest groups to see if something new (either a new signifying object, or new annotations on an object) is added. The

user can just sit back and digest the content. Much of the content 'fed' to the user is pulled from profile pages of her connections. However, the feed can also contain ads, or even signifying objects about people or groups to which the user herself is not connected. I will get back to this in section 5.6.

The content displayed in the feed is determined by the medium's settings and algorithms (Gillespie, 2014, p. 167). Especially algorithms play an important role in the order and selection of the displayed content: they select, include, exclude and rank the content that is displayed. Algorithms are "encoded procedures for transforming input data into a desired output. The procedures name both a problem and the steps by which it should be solved" (Gillespie, 2014, p. 167).[17] The design of algorithms rely greatly on their developer's knowledge, limitations, expertise, and choices (Kitchin, 2017, p. 18). The result is that the values and views of the developers are (consciously or unconsciously) scripted into the algorithm (Mittelstadt *et al.*, 2016, p. 1). Meanwhile, the parameters of the algorithm often "contextually weighted and fluid" (Kitchin, 2017, p. 21).

A prominent example of a feed algorithm, is Facebook's 'EdgeRank'. As Kincaid explains, the selection of the displayed content by EdgeRank is based on the calculation of three main factors: (1) affinity, (2) weight, and (3) time.[18] 'Affinity' is the score attributed to the relation between the user and the signifying object's publisher. This score is calculated based on the frequency and type of interaction between the user and the publisher. The 'weight' concerns the attention value given to the signifying object by users: the more users comment on and react to the object, the more weight it gains. As such, the 'weight' can be seen as the 'attention value' of an object. In this sense, the annotations express a certain materialised attention. Lastly, the factor 'time' adds a decay rate to the signifying object: the more time passes the less important the signifying object is considered to be. These factors combined lead to a certain score for the signifying object. The higher the score of the signifying object, the more likely it is that the signifying object will be displayed in a user's News Feed. Additionally, with changes to the user's profile, her feed is adjusted to parameters that fit her behavioural communicative pattern (Bucher, 2012; Beer, 2009).

With factors like 'affinity' and 'weight', feeds based on algorithms like EdgeRank are inclined to display a circular logic (Bucher, 2012, p. 1169): as the feed imbues its displayed signifying objects with a certain prominence and pretence of importance, this boosted presence has a bigger chance to trigger more user

[17] There is no consensus on what an 'algorithm' exactly is (Hill, 2016, p. 37). The use of the term 'algorithm' in the public discourse is often broader than what an algorithm strictly speaking entails as people also use the term to refer to larger assemblies (Mittelstadt *et al.*, 2016, p. 2). Hill offers a more precisely formulated description of algorithms by arguing for the definition of an algorithm as "a finite, abstract, compound control structure, imperatively given, accomplishing a given purpose under given provisions" (Hill, 2016, p. 47). Because reproducing Hill's argumentation runs outside the scope of this study, Gillespie's description will suffice. I would like to refer readers interested in a thorough discussion of what an algorithm is to Hill's paper *What an Algorithm Is* (Hill, 2016).

[18] Jason Kincaid, "EdgeRank: The Secret Sauce That Makes Facebook's News Feed Tick", *TechCrunch*, 2010. `https://techcrunch.com/2010/04/22/facebook-edgerank/`, last accessed 2017-05-04.

attention, which in turn may reinforce the object's visibility. By shaping the primary information flow in feeds, social media thus not only increase the presence of particular references, but can even magnify their presence in a self-affirming cycle of attributed importance. As such, the feed reinforces the visibility of certain connections and signifying objects, while rendering others less visible, or even invisible in the feed (Bucher, 2012, p. 1169-1171).

Feeds display a relatively strong expression of technological intentionality: they aggregate, rank and display without needing any human intervention. With their emphasis on objects that invoke user reactions and active user relations, feeds normally prioritise signifying objects that generate users' reactions. By placing the most engaging content at the top, the feed provides an incentive to users to also participate and communicate (Bucher, 2012, p. 1175). With this focus on the actuality and/or popularity of signifying objects, combined with the affinity between users, the quality of the content seems to have little influence, except indirectly if it evokes user interaction. Feeds construct "a forum that makes the everyday newsworthy" (Ridenour, 2011).

While some applications give users some control over the feeds by providing them with options to shape (part of) the feed to the user's own taste (e.g., by allowing the user to include or exclude certain connections from the user's feed)[19], the medium controller is in charge of the design of the feed and decides which options are made available to users. As such, they have a significant influence on what is presented to users in their feeds. The medium controller's perception of what is 'valuable' — and hence what feeds should prioritise — therefore plays a crucial role. The difficulty here, is that users and the medium controller are likely to hold different views on what is valuable content. Users generally consider a certain signifying object intrinsically valuable due to the information that it provides them (whether it be informative, entertaining, has personal value, etc.). As such, the information has a certain use value for the users, and the users *therefore* spend attention on the object. Media controllers, on the other hand, find certain signifying objects valuable *because* users spend attention on it. They can use the information to tweak advertisement processes, or sell the information to third parties. The information has *exchange value* for the medium controller. Medium controllers therefore likely favour an information flow and signifying objects that invite high user interaction and attention time (or return visits). Their interest lies in the quantity of content and the frequent interaction of users over the quality of content (van Dijck, 2013, p. 161). As such, the feed embodies an information flow — and meaning — that is managed by an infrastructure oriented on the exchange value of information, while the flow itself serves users for whom the value lies in the content. The consequence is that the user's cognition is occupied by an information flow that follows a non-informative principle: "the principle of economic competition, the principle of maximum development" (Berardi, 2009a, p. 37). However, the mechanisms that embody this friction are hidden from view for users: users encounter only the end-result in a smooth interface and

[19] See "Controlling What You See in News Feed", `https://www.facebook.com/help/335291769884272?helpref=faq_content`, last accessed 30-10-2019

are given little insight into how the algorithms work and how they weigh various factors. Users may even be unaware the information flow is controlled for them by means of algorithms (Ippolita, 2015; Eslami *et al.*, 2015). The feed therefore puts users at risk to assume that the information that is presented to them is also the most valuable information for them, while in fact the information that they themselves may find valuable — either to receive, or to transmit to others — could be underexposed.

5.5.3 Presence over time

The social media architecture influences how signifying objects are present on the platform over time. In order to guarantee ongoing user input, users are continuously invited to provide new information by adding status updates about what they are doing or thinking, while the feeds prioritise new content by placing the new at the top of the page. And it is not just the feeds that focus on the new: also the walls of the profile pages and group pages generally have their publications presented in a chronological order with the newest at the top. As such, social media can encourage actuality to the extent that "individual moments transform into overall flow—a *feed* of now now now" (Bogost, 2010, p. 28).

Despite this focus on the now, the past does not disappear on social media: the signifying objects are generally stored by default. Depending on the user's settings, her profile page can display a collected past of every post she or others published on her page since the creation of her profile. If the user takes no action, the traces of social interaction are indefinitely retained — and accessible (that is, unless the media controller decides to remove a signifying object at the request of another user or because the content is in breach with policy). If a user wants to delete past signifying objects or restrict the access of audiences to them, the user will need to do this manually. Because this is time-consuming, users are likely inclined to let the signifying objects remain on the system. The consequence is that their social communicative interactions often remain long-term accessible to their audiences in this digital tertiary memory and thereby gain a certain persistence (boyd, 2010, p. 46). Tredinnick argues that this leads to the construction of a digital archive that "saturates the entire social network. By sharing information we become part of a living archive" (Tredinnick, 2008, p. 164-165). In this context, we can see the user profile page as a personal archive that records expressions of the user and the responses of others to her. However, it is important to note that this is indeed a 'living archive': as social media utilise the flexible affordances of digital information by allowing users to edit, delete or hide certain content, the past on social media can always be adjusted, rewritten or deleted by the users involved. The editing and deletion of signifying objects can decontextualise content or reorganise a historical view on events, thereby affecting the content's meaning. The flexibility of the signifying objects thus forms a risk to the authenticity of online expressions, especially because they be changed without leaving a trace (boyd, 2010, p. 54).[20] Facebook even offers the editing option 'Change Date',

[20] From the perspective of social media platforms editing and removal of content may actually

which allows users to change the date of content (although a user can only change the date to a date previous to the original date of the publication, not after), thereby affecting its positioning on the profile page. Especially in the case of a thread containing a discussion, the editing or deletion of comments can mangle the context of the comments. The consequence of all the editing options is that while much of the content may remain available over time, there is no guarantee that it is an accurate reflection of the past.

Additionally, due to the editing and annotation options, the past can have its presence revived. When users annotate a particular object, they increase its (attention) weight and potentially push it to the top of the feed, thereby extending or even reviving the reference's presence into the 'now'. For instance, if a curious user decides to scroll through old publications of a friend and annotates them with for example a comment or a 'like', these publications receive new weight, which may move them back into the feed and under the attention of other users, where it in turn may trigger more reactions, and so on. With annotations, old signifying objects can therefore easily receive a second round of attention.

While old content remains and may easily go up for another round of attention, it can be difficult for users to retrieve a *specific* old object (Lovink, 2016, p. 31). Older material is less present and search features on social media tend to be poorly equipped for the task. Users will often need to manually locate a signifying object by scrolling through personal profile pages. Lovink therefore argues that social media are not moving in the direction of allowing users to remember everything. Instead, "[a]s only temporary reference and update systems, difficult to access with search engines, the streaming databases are caught in the Eternal Now of the Self" (Lovink, 2016, p. 31).

5.6 Connected publics

Social media mediate between users and their audiences. Due to their technological character they afford communication between more people at the same time than was previously possible (Cuonzo, 2010, p. 174). As with basic websites (see section 4.5), the size of the potential audience is less constrained by space than any offline setting, while the access time is often stretched longer than most offline interactions (face-to-face communications, telephone conversations, etc.) — if there is an end to the access time at all. However, the manner in which social media construe audiences is fundamentally different from basic websites: audiences on social media generally are *connected* audiences. Most social media offer their users a feature to 'connect' to others. This allows users to establish a certain relation between them and others, whether it be other users, interest groups, companies or governmental institutions. These connections that users establish, form the core of the audience composition on social media.

Connecting is a mix of a technological feature and a social choice. The

be beneficial for profit purposes, because it entails more user activity as well as a glimpse into user behaviour.

user therefore plays a pivotal role in the composition of her audiences. A user commonly connects to others by finding their profile on the platform (by means of searching their names or email addresses), or by suggestions made by the platform architecture. Usually, users connect based on shared interests (boyd, 2010, p. 45). However, these shared interests can diverge substantially, leading users to connect based on a mix of various types of social relations (Trottier, 2011). Users even connect for diplomatic reasons to people that they do not like (Meikle, 2010, p. 18). The resulting audience is generally comprised of a mix of others that have 'strong ties' to the user, as well as those that have 'weak ties' to her (Nahon & Hemsley, 2013, p. 31-32). When users have strong ties, it means that they have a strong information flow between them, tend to be like-minded and have a high degree of overlap in their networks (Nahon & Hemsley, 2013, p. 31). The strong ties are therefore generally the people to which the user is 'close' also in the offline world (friends, family, etc.). However, by generating a field for undemanding and open interaction (see section 5.4), social media make it attractive for users to expand their connective network to 'weak ties'. Weak ties are often others that are embedded in a different social context, for example, they do not live in the spatial proximity of the user, or do not share certain characteristics like language, common-experience, age, etc. (de Meo *et al.*, 2012). While users do not regularly associate with their weak ties in offline or directed settings, like by means of email or the telephone, they do tend to connect to these weak ties on social media. This even happens to such an extent that that most connections on social media like Facebook tends to consist of weak ties (de Meo *et al.*, 2012). By connecting to weak ties, social media set up an information flow that would not have existed without this mediation, as they grant distant others a continuous insight in the user's life (Ridenour, 2011). Combined with the open and undemanding character of this information flow, "social networking empowers acquaintances to contribute to our lives in ways previously reserved only for friends" (Hamington, 2010, p. 142). Because connecting is the pillar of social media use, many social media actively suggest potential interesting connections to users and "promote connectedness as a social value" (van Dijck, 2013, p. 11). A user's connectivity may be further triggered by the 'popularity principle', which means that "the more contacts you have and make, the more valuable you become, because more people think you are popular and hence want to connect with you" (van Dijck, 2013, p. 13).

Connecting is generally not a one-sided act: users need to mutually agree to connect to each other. Because the user profile is the anchor of the connection, it plays an important role in the choice of users to accept or decline a connection. If people do not recognise or acknowledge a particular user as a party to which they want to connect, or believe being associated with her may harm their self-presentation or reputation, the user may miss out on social connections. This may explain why users are inclined to use avatars that portray a recognizable photo of themselves or something closely tied to their real life persona, like in the case of mothers using a photo of their children (Wittkower, 2014); revealing something fundamental about their identity may make them more easily identifiable for others who may know them offline and want to connect, or invite unknown others to

connect because they are drawn to the user's self-presentation. However, giving shape to a profile can be challenging if the user wants to connect to different kinds of audiences, because the user profile commonly is a single baseline profile for all audiences (van den Berg & Leenes, 2010; Wittkower, 2014). In the case of Facebook, the name, profile picture, banner and biography is the same for all the audiences that have access to the profile (see figure 5.2.1). This requires users to construct a 'one-size-fits-all' proxy for possibly highly diverging audiences. Any failure on this level can lead to the sharing of information with certain unintended audiences or to pushing particular users away. For example, take a referent that is known to one group of friends under a certain nickname (e.g., because she knows them from playing World of Warcraft, where people play under a character name), while another group only knows her under her 'real life' name. The referent may then experience a difficulty when constructing her basic profile, because she will have to choose one particular name that she to identify herself to multiple audiences and be recognised by them. The referent could solve this by using both names in one, e.g., "Paulan 'Caligari' Korenhof", but then she immediately collapses her two different roles for her audiences. Building a profile that can be used to connect to different publics can therefore be challenging.

Once a connection request is accepted, the parties become technologically linked in the platform's database and are set up for certain access authorisations with which they generally get (increased) access to each other's personal information. As such, connecting, while being on the front end level a social act, is in the social media architecture a technological action that ties the presence of signifying objects to the technological individuation of users. The act of connecting and its implications for content-access therefore has a deeply technological nature. An example that shows this, is the effect that the 'tagging' of referents in pictures and other posts has on the access of the referent's audiences to the content. When users 'tag' someone in a signifying object, this act does not only establish a reference by adding the name of a referent to the content, but it also establishes a link in the database between the referent's profile, her connections and the signifying object. This link then provides the connections of the 'tagged' user access permission to the object.

However, the establishment of a connection between users is not always necessary to access the objects, nor does it necessarily result in a full access of the connection to all the content relating to a particular user. There can be many differentiations with regard to information access on social media. Contrary to the baseline profile which generally is presented equally to various audiences, social media like Facebook allow a user to select a particular audience for each object that she posts on her wall (see figure 5.2.1). For example, she can make a signifying object accessible only to specific connections, to all connections, to connections of connections, or she can make the content accessible to all audiences. There is a high degree of variety in these options, and the potential audience of signifying objects can range from one specific person to a worldwide audience.

While these diverse options are offered, using them to properly frame the complex nuances of human social relations is difficult; the user needs to divide

feelings and relations into technologically articulated categories. For example, on Facebook users can choose to manually differentiate between various types of relations by splitting their 'friends' in distinct subgroups like 'family' or 'colleagues'. However such differentiations are generally still a superficial subset of offline relations (Leenes, 2009, p. 57). This results, at least partially, from the fact that social media interfaces tend to be poorly equipped to deal such different sorts of relations (Losh, 2010; van den Berg & Leenes, 2011). The technological mediation forces users to reorganise the hierarchy of social relations (the different information sharing relations that people have with intimate friends, close friends, friends, acquaintances, etc.) that people generally maintain in offline situations, into the relative rigid connective structure of the platform. The consequence is that, in the end, connections are often treated equally on social media, thereby leading to a relatively 'flat' framing of social relations in which distinct relations (e.g., close friends and colleagues) receive equal access to personal signifying objects. When publishing information on social media users thus often publish for a heterogeneous audience, this while they likely have a specific audience in mind (e.g., hobby related, friends, professional) (Meikle, 2010, p. 14).

The social media architecture (and especially its default settings, as users tend to accept these) thus highly affects the audience that can retrieve the content. As the social connection between two people is translated into a piece of code in a database, the architecture presses its own logic on the human relation and corresponding audience composition. A good example of this is the Facebook's 'friends of friends' setting. With this setting, a user makes her content available to all her connections, as well as to all the connections of her connections. With this setting, Facebook "transforms a discrete set of users into an audience—it is a group that did not exist until that moment, and only Facebook knows its precise membership" (Gillespie, 2014, p. 188). Such settings can easily make the content accessible for a massive audience of which the exact composition is unknown to the user. However, this setting seems currently to be little used (in a 2017 study, only 6% of the users reported to use this setting Fiesler *et al.*, 2017).

What further complicates the user's overview of her audience, is that the social media interface generally hides the way in which the technology establishes a relation between a user and her audiences. Social media users can therefore easily misunderstand or misjudge the character and scope of their audience (Leenes, 2009; boyd, 2010; van den Berg & Leenes, 2010; Wittkower, 2014; DeVito *et al.*, 2017). Users may therefore easily err in properly segregating their audiences. Especially with regard to signifying objects that are accessible to mixed connections like a thread, getting an overview is a difficult task due to the role played by various personal privacy settings of the connections. One of the major risks for social media users is therefore that their publication reaches an unintended audience. I will discuss this in more detail in the final section of this chapter. Even in the case of social media aimed at public communication, such as Twitter, it is questionable whether users and audience members truly grasp what 'publicly accessible' de facto means in the online realm, because a tweet can freely reach various cultural contexts and invoke different interpretations of the signifying object.

Lastly, it is important to discuss another level on which social media affect the relation between the user and her audience. Although social media thrive on interaction and promote activity by inviting the audience to participate by offering them easy to encode reactions, the invisibility of the audience combined with the broadcasting character of the content relieves the audience from social pressure to react. As such, social media can easily give rise to voyeurism, snooping, and even stalking (cf. Hill, 2009; Leenes, 2009; Lyndon *et al.*, 2011). While many users will be aware of this, this awareness can have a side effect: "The potential of being watched by others contextualises their own surveillance. Not only does this suggest that surveillance is rampant on the site, but it also dampens users' ethical concerns about covertly watching others" (Trottier, 2011). The consequence is that on social media we see the rise of a new dimension in the relation between the user and her friends, family, colleagues, acquaintances, etc.: "Mere contemplation and passive observation have replaced actual communication, and social relations seem to become—to some extent at least—merely *looking* at other people, transforming our friends, in our eyes, from active participating subjects into objects of interest and entertainment" (Vejby & Wittkower, 2010, p. 102). Especially with the help of feeds, the audience members can conveniently consume these signifying objects from one location by merely scrolling up and down. By doing so, the feeds efficiently provide entertainment and gratify the socially curious nature of human beings (cf. Fairweather & Halpern, 2010).[21] By enhancing or reducing the presence of publications by others, the feed heavily affects the "the relationships users are encouraged to maintain" (Mittelstadt *et al.*, 2016, p. 10). Poorly connected, uninteresting or unpopular users will hardly be visible to an audience due to their meager ranking on feeds. The social media architecture, and especially the feed, thus heavily affects the actual audience as well as the relation between the audience and the publisher.

5.7 Complications of the presented persona

Social media are generally used for social interaction and self-presentation. The medium depends on its users for content. However, the architecture of the medium also impacts the presence of personal information, the content that is encoded, as well as the manner in which the informational persona can be 'compiled' for the perception of its users. In this section, I will combine the findings of this chapter and discuss how they together affect the formation of the informational persona, and how this may represent a referent in a problematic manner.

On social media, the informational persona is constructed at the axis of user activity and user connectivity. Users interact with others in a framework that generally aims (in accordance with the medium controller's interests) to advance this interaction. Social media therefore tend to promote a culture where 'sharing

[21] Despite the fact that the implementation of feeds can encounter some initial resistance by users, users tend to accept them eventually (Trottier, 2011).

is caring'[22], thereby framing the revealing of information as social activity. As such, social media invite their users to be 'hyper-expressive' (Berardi, 2009b, p. 180). Users generate the content, and are encouraged to engage with content of others. The potential problematic aspect of the content is tied to the character of social media; promoted as medium for *social interaction*, it invites users to use the medium to connect to others, share personal stories, photos or anecdotes with others, while the platform itself tends to publish on the informational acts of the user. This personal level of the content necessarily impacts the informational persona: the content is highly personal.

Also, the platform itself can add content to the informational persona, like the number and identity of connections of a particular user (although sometimes this feature can be turned off). By making such new types of information visible, the platform allows users to see themselves as well as others in new contexts and affect their interpretation of the referent. For example, based on the number of connections, people could conclude that the user is popular, an attention seeker, interesting, not socially skilled, unfriendly, privacy aware, etc. Also, the quantification and visibility of the connections can affect the self-perception of users: they may conclude that they have a lot of friends, are popular or unpopular, or maybe just that they have a lot of empty relationships. The explicit visibility of social connections can make a person's relations a more prominent ground for discrimination (cf. boyd, 2014), or can complicate getting a job.[23]

Next to, but also tied to, the highly personal level of the content, social media impact the content of the informational persona at another important level: the identifiability of the referent. Due to the social and personal level of the interactions on social media, combined with the policies and affordances of the platform, we see that social media users often reveal themselves — at least to some degree — as an identifiable offline person. Users easily leave many traces to their offline lives, even if they do not use their real names. For example, their connections, pages that they are interested in, events they sign up for though the platform, can all reveal the user as a particular offline individual. On social media, the ties between the online and the offline are relatively strong: "offline contexts permeate online activities, and online activities bleed endlessly back to reshape what happens offline" (Baym & boyd, 2012, p. 327). The presence of the informational persona on social media therefore entails relatively high risks for the offline individual.

Moreover, we also see the integration of commercial references in the informational persona as users add content to their profiles in response to the promise of a prize if you 'like' a certain product. Besides the integration of commercial elements by users themselves, third parties can also themselves collect and 'repurpose' user content for advertisement goals. This can heavily affect the presentation

[22] As is nicely portrayed and pushed further by Dave Eggers in his book *The Circle* (Eggers, 2013).

[23] Sarah Quinn, "Facebook costing 16-34s jobs in tough economic climate", *On Device Research*, 2013. `https://ondeviceresearch.com/blog/facebook-costing-16-34s-jobs-in-tough-economic-climate`, last accessed 05-09-2018.

of personal information: the subject's reference is processed into a new context with which she may have no connection at all. This can give rise to false references that can lead users to misguided interpretations of the referent, or even be hurtful to the referent or her social environment. A painful example of this was when a third party company scraped a photo from the Facebook profile of a 17-year old rape victim who committed suicide after being cyberbullied, and used it for a dating ad with the text "Find Love In Canada! Meet Canadian girls and women for friendship, dating or relationships! Signup now!"[24]

Meanwhile, once a signifying object is encoded, it can become an object of interaction: the social media audience is an active audience that (if it does not retreat into the role of voyeur) can take on the role of co-publisher by annotating the user's publications. By doing so, the audience affects the content and the meaning that is given to it (de Fina, 2016; DeVito *et al.*, 2017). In turn, the annotations of others on the medium, tell the referent what they find important or noteworthy about her.

On social media, we thus see the rise of an highly personal informational persona of which the content is shaped in an interplay between the initial sender, other users including potential third parties that annotate the sender's content, and the social medium. The construction of the persona thus takes shape in a 'triad' intentionality in which the users themselves play the most decisive role: without their encoding of new content, annotations, and establishing connections, nothing happens. This triad intentionality also affects the presence of the persona, but with a different division of roles. A view on this persona is presented in two distinct situations: (1) on a profile page, and (2) in the feed. I will first discuss the general presence of the informational persona on the profile pages, before I go into the presence of the persona as established in feeds.

The core of the informational persona on social media is rooted in the user profile. Here, most of the content related to a particular referent is collected. By requiring the creation of a profile in a certain baseline format and giving it a central role, we see a relatively strong expression of the intentionality of the social media's technological architecture on the formation of the persona. On social media the user is represented *as* an informational persona: the user has to present herself in an one-size-fits-all-audiences identity frame, while she is individually highlighted in her every activity on the platform, her profile identity accompanying her every comment, with all her information accumulated on her profile page. On social media, users are thus constantly working on redefining themselves as they add content in an interplay with others and the mediating technology. The triad intentionality which gives shape to the presence of the informational persona on the profile page accommodates the interests of users, as well as of the medium controller. In this, the mediating nature of social media often harbours two distinct interests with regard to the information flow in the medium; while the users are interested in social interaction and the content of information, the medium

[24]Helen A.S. Popkin, "Bullied dead girl's image used in dating ad on Facebook", *NBC News*, 2013. `https://www.nbcnews.com/technolog/bullied-dead-girls-image-used-dating-ad-facebook-4B11187466`, last accessed 20-01-2019.

controller's interest generally lies in running a business. The self-presentation of users on social media, entails therefore also at the same time a commodification of their persona. The mix of self-presentation and commodification give rise to a particular presence of the persona: by encoding their selves into the social media architecture, their personal information is materialised as a product and subjected to an architectural regime that subtracts exchange value from it and offers it to audiences often looking for entertainment. As such, the user is submitted to "a form of subjectivation that is both infiltrative and extroversive" (Dyer-Witheford, 2015, p. 93). The persona is presented as an ongoing list of objects that represent her in certain settings and contexts. As the user is constant invited to add updates by the eternal "What's on your mind?" status bar, the persona can end up being a detailed portrayal of the subject's life, moment-by-moment. Here, Debord's *Society of the Spectacle*, first published in 1967, seems to be a premonition for a praxis perfected in social media when he states: "In societies where modern conditions of production prevail, all of life presents itself as an immense accumulation of spectacles. Everything that was directly lived has moved away into a representation" (Debord, 1977, §1). With this, Debord predicted the rise of economic models that pushed 'appearing' as main value towards the foreground of social life, instead of the classic value of 'having' (Debord, 1977). In order to 'appear', users need to encode their selves in a prefabricated uniform format that allows the media controller to commodify this information. With this, social media platforms are turning the informational persona into a 'spectacle' (cf. Baroncelli & Freitas, 2011; Vejby & Wittkower, 2010; Virno, 2003). In the spectacle, "human subjects find themselves faced with objective forms that they have themselves created, into which they have alienated their own attributes and capacities, and which, despite being expressions of their own selves, appear to be quite independent and separate from them" (Bunyard, 2017, p. 18).

The presence of the persona in the feeds is an even stronger reflection of a spectacular presentation of the persona than the profile page. In the feed, users are generally represented by one or a few actual and popular references, woven in between others. As the presence of information is tied to the actuality and popularity of signifying objects, references to big life events and trivial ramblings may gain an equal status in the portrayal of the persona. The users themselves have little to no control over the manner in which they are represented in the feed. This complicates the user's self-presentation, because she has little means to emphasize to others what she feels is really important or defining of her by increasing the presence of a particular reference. She can try to increase the reference's presence by for example repeating the same status update over and over or by linking and commenting on the post herself, but this in turn may reflect her negatively as others could interpret this as for example neurotic or narcissistic behaviour. The result is that the defining traits of the persona as shaped by the presence of references is highly dependent not only on the user herself, but also on the social media architecture and the actions of the user's connections. The presence of the persona in the feeds thus also takes shape in a 'triad' intentionality, but this time with a role division in which the technological intentionality has the

most weight.

The persona in the feed is constantly on the move in a cyclic manner; a new perspective on the persona is peaking with every signifying object that makes it to the top of the page, either by being new or by receiving new annotations, then the particular perspective decreases from there on, and peaks in a new form with the next object that is actualised in the now. With the social media's general focus on the 'now', the problems of the informational persona are overall of an *immediate* nature. However, as stated, the past can unexpectedly rear its head and be revived by others who comment upon it. The risk of older content on social media is that it may be edited, parts of threads and the like may be deleted, or may have lost the connection to its context. As the existence of expressions can stretch far beyond the time frame of the interaction, the context of ad hoc communication and discussions may easily erode and the content "may lose its essence when consumed outside of the context in which it was created" (boyd, 2010, p. 46). Currently, the burden lies with the user to prevent any decontextualisation of the past; it is the user herself who can delete old signifying objects (which may in turn lead to the decontextualisation of connected other content) — with due note that the tools social media offer for this are burdensome. Initiatives to ease this burden of manually deleting old content, like the 'Web 2.0 suicide machine'[25] are not always (or generally not) welcomed by social media controllers and are blocked from use.[26] The reins of the engineering of sociality thus generally remain firmly in the hands of the medium controller.

The view that a particular user has on the persona thus takes form in an interplay between new, but uninteresting, signifying objects which quickly lose visibility, and old, but interesting, (sometimes unforeseeable) popular and possible decontextualised objects. In this manner, the referent is represented to others and herself as her latest fling or by that which evokes reactions of others. The more others react to a particular signifying object and share it, the stronger and more persistent the presence of its reference becomes. This can increase to the point that the reference goes viral, which I will discuss in chapter 7. The cyclic presentation based on actuality and activity is likely to entail a relatively superficial portrayal of the referent that does not seem to do justice to humans as beings with a history and a variety of life experiences. However, the quick and easy consumable character of content is part of the attractiveness of social media. In order to be part of the social interaction and strengthen the relation to their social connections, users need to display themselves and engage in this cycle of recurring presence by continuing to publish and react in order to maintain a certain relevance and receive social gratification. On social media, people therefore practice the art of socially appearing — but appearing to whom?

The 'who' that forms the audience of content is one of the potential biggest complications of the informational persona as portrayed by social media — both

[25] `http://suicidemachine.org/`, last accessed 30-10-2019.

[26] See e.g., Paul McNamara, "Facebook blocks 'Web 2.0 Suicide Machine'", *Computerworld*, 2010. `https://www.computerworld.com/article/2522527/facebook-blocks--web-2-0-suicide-machine-.html`, last accessed 03-07-2019.

of the persona as presented in the feed, as well as presented on the profile page. First of all, there is a friction between the 'social' on social media, and the 'social' in the offline world. As discussed in section 5.6, the complexities and vast array of nuances in human relations are generally poorly reflected by the mechanisms of the social media architecture, while these settings are difficult to operate. Secondly, on social media users generally have a poor overview of their actual audiences. They need to deal with an invisible audience of which the composition is opaque and often influenced by factors invisible to the users. Especially the feed can cause trouble on this front: the technological intentionality of the feed highly impacts the actual audiences of content, while it gives the user herself little clue as to how, when, and to whom her content is presented. Here, it is important to note that the accessibility of the content does not guarantee an audience (boyd, 2010, p. 48). While users have some control over their 'non' audiences, i.e., the audiences that (initially) do not have access to the content because the user restricts their access by means of the medium's privacy settings, users thus have little to no control over the audiences that actually are confronted with the content. The user's persona may thus have anything from a very strong to a very weak presence for certain of her connections — but she does not know what and for whom. She therefore has a poor view on how she is represented to others. Her audience may consist of many connections that mean little to her or for whom she in fact does not want to have a strong presence. The user may even have little to no actual audience, or lacks audience members that are vital to her, and be unaware of this. As identity construction takes place reflexively in interaction with others, the lack of response of pivotal others, or the added response of undesired others, can affect the user's identity construction in unwanted manners.

The lack of refined audience composition settings combined with the lack of a good overview on her audiences, poses the self-presenting user a challenge. To refer back to Goffman's research on self-presentation as discussed in section 2.3, in order to perform convincingly in different social roles and maintain distinct social relationships, a user will need to give clear signs about her role to her audience, while keeping personal information that does not fit this role 'backstage'. One of the main strategies that people employ in order to successfully do this, is to segregate their audiences based on the role that they aim to play for that particular audience (Goffman, 1959, p. 137). For example, people tend to reveal different things about themselves to close friends in an intimate setting, than they do to students when giving a lecture. The segregation of these audiences is often tied to a particular region, e.g., a classroom, a bedroom, a bar. On social media, the character of the space in not clear, and can therefore give rise to a 'regional ambivalence' due to which the user misjudges the context in which she is interacting (Wittkower, 2014). As the user acts in a ambivalent and opaque technological space with a poor overview over her audiences, it can be difficult for her to take on different social roles and maintain certain distinctions in her relationships. Lack of audience segregation is therefore "one of the most prominent issues of social software" (Leenes, 2009, p. 48) — and a task which is contradictory to some of the mechanisms of the medium's architecture (e.g., features like the

'share' button, the 'friends-of-friends' setting, automatic publishing, the single profile). If the medium's audiences exceed the user's expectations, they can compromise the contextual integrity of the user's publication (cf. Nissenbaum, 2010). As a result, different, previously separated, social contexts may collapse and the user may be sharing her information with a bigger group of people than she realises. Such collapses can complicate and even disrupt the user's self presentation. And although the collapse of social context is in itself nothing new, the digital affordances and infrastructure of social media are likely to entail an amplification of the scope and intensity of such a collapse: the audience is easily much bigger than in offline situations. Additionally, due to the lack of transparency of the audience, the user may be unaware of any audience segregation failures and thereby miss out on opportunities to adjust her performance or repair the damage (boyd, 2010, p. 50). Without proper audience segregation options, online users may need to present themselves as 'flat characters', so that their informational persona is suitable to a wide array of audiences (Leenes, 2009). The difficulty of social media lies thus not in the fact that an audience *is watching* them, but in the *scope and composition of the audience*.

An additional issue is that the audience itself can lack a good overview of the content and its (original) context and may thereby misinterpret the referent. As social media utilise the flexible affordances of digital information by allowing users to edit, delete or hide certain content, the audience can easily be confronted with content that over time is adjusted, rewritten or has parts of it deleted. This affects the context of signifying objects, especially in threads and timelines: the editing and deletion of signifying objects can decontextualise and reorganise their as well as surrounding content, thereby affecting the meaning of a comment in ways potentially unintended by the expresser. Especially in the case of a thread containing a discussion, the editing or deletion of comments can mangle the context of the remaining comments. The consequence is that the audience, while often being able to access older content, may miss vital parts of the original context and may misinterpret the meaning of the content — this in turn may reflect problematic on the referent.

This brings me to the next point: these others can themselves also be the cause of difficulties. On social media, the other is the reason to be on the platform, but at the same time she is also a voyeur, an accomplice in the media's architecture, and a liability. While the technological mediation invites users to watch others, share information and abide by certain norms, it is in the end the user who decides to do this. User practices and norms therefore play a crucial role in the shaping of profiles and the construction of online identity. For example, on a medium where the user norm is to use a nickname, new users will be inclined to follow this norm. This brings me to what I take to be one of the most problematic issues of the informational persona on social media: the commonplace practice to publish or forward personal information about others and identify others by means of 'tags' and the like (see section 5.4.3). Boosted by the architecture of social media, the audience is always also a potential publisher and can easily reveal information published for a selected audience to other audiences, or disrupt the presentation of

personal information by scattering the context. The other does not have to have bad intentions — even good intentions can affect the interpretation of someone's informational persona negatively. To give an anecdotal example:

a friend knows me for loving Hammer horror films from the 1960s. When one day she came across a discounted film box which she thought was the same genre, she posted the link to the box on my Facebook wall. The link was in fact for a box set of 1970s nazi sexploitation. Needless to say that they are not the same, and I deleted the post with lightning speed — hoping that no one has seen it and would associate it with me.

Even when a user manages to run a smooth audience segregation on social media, and publishes with discretion, her informational persona can easily be affected or spread in a negative or unwanted manner by others. If these others have a higher connectivity than the referent, they are able to generate references to her with a stronger presence than the referent can do herself. The consequence is that others may have a stronger impact in shaping an informational persona than the referent herself does. As such, social media manage to give a new dimension to Sartre's expression "l'enfer, c'est les autres" (Sartre, 1987).

The question is whether art. 17 GDPR can resolve the issues identified in this chapter. I will discuss this in chapter 9.

Chapter 6

Search engines

Contents

6.1 Introduction

Did you Google him?
by singlethirtysomething

(. . .)Now I know I practically live online, but really, if you're meeting someone new, it's common sense to check him/her out by doing an internet search. (. . .)

CS said she'd met up with a guy who had Googled her and she was really disconcerted that he knew things about her that she hadn't told him. My take is that it showed he'd been interested and done his homework. If you're honest with each other, all that information will come out eventually anyway, so why hide it? CS reckons she can learn all she needs to know about someone within the first 10 minutes of meeting them. Personally, I reckon my instincts could use some factual backup.

It is vital to do things like apply high privacy settings (. . .) and control what sort of information you allow to be in the public domain. Not only does this protect you personally but it also prevents your boss seeing what you were up to at that tequila night last weekend. . .

Do you Google your dates? Is it intrusive or sensible?

singlethirtysomething[1]

The text above is an excerpt from a blog written by 'singlethirtysomething'. She describes a discussion between friends about the use of the Web for their dating activities. In this, she ascribes a key role to online search engines.

The digitisation of information afforded new retrieval mechanisms that shape how the information is revealed (Nissenbaum, 2010, p. 56). Search engines are an example par excellence of such digital retrieval mechanisms. Search engines have been part of the Web in different shapes and forms since roughly its beginning — their roots even go as far back as information retrieval research in the 1960s (Hendler *et al.*, 2008, p. 62). Online search engines are built to serve the purpose of mass use information retrieval. Examples of current popular search engines are 'Google Search', 'Bing', and 'Duckduck.go'. Search engines are immensely popular and incorporated in the regular practices of many Web users (as in the case of 'singlethirtysomething') and are "crucial in connecting audiences to content" (van

[1] Singlethirtysomething, 2009. `https://singlethirtysomething.wordpress.com/2009/01/14/did-you-google-him/`, last accessed 04-03-2017.

Dijck, 2013, p. 121). Whenever we do not know the URL of a certain website, or even what website we should or could be looking for, the little magnifying glass in the top corner of many Web browsers offers us a solution — or at least a starting point — for our journey into the online world. After entering a search term, the user is offered a search result list that displays the potential websites matching her term. The search term can be anything, including individual names. And as the blog above shows, the use of search engines to locate personal information is not an uncommon practice. Yet, despite — or because of — arguing in favour of the use of search engines to look up information on potential dates, the author concludes her blog by urging readers to restrict the accessibility to their personal information in the public domain.

The display of search results containing personal information was at the heart of the dispute in the heavily debated *Google Spain* case.[2] In this case, a Spanish citizen wanted to have two search results erased that were returned in response to a search on his name. The search results pointed to two small newspaper articles from 1998 stored in the archive of the newspaper *La Vanguardia*. The articles contained information about the forced sale of the subject's house as a result of social security debts.

In May 2014 the CJEU ruled in favor of the subject. It stated that a search engine can be required to remove search results if the content to which they refer has lost its relevance and is a disproportionate burden for the individual.[3] What made the case particularly interesting, is that the court case focused on the responsibility of the search engine provider as a technological driven intermediary, and not on the original publisher of information. The case gave rise to a broad interdisciplinary, but also polarised, discussion about the impact of search engines as a mediator of information. Exemplary for the discussion are the contrary views of the Advocate General Jääskinen who advised the CJEU, and the CJEU itself.

According to Jääskinen, a search engine is a passive mediator that provides a truthful reflection of relevant web pages to the users.[4] Jääskinen argues that as mediator, a search engine 'only indicates' where on the Web a user can find already existing content that is made available by other parties.[5]

Contrary to Jääskinen, the CJEU views the actions of a search engine as "*additional* to that carried out by publishers of websites" [my emphasis].[6] In the ruling, the CJEU argues that the mediation of the search engine impacts the presence of personal information beyond the source websites because the search engine "enables any internet user to obtain through the list of results a structured overview of the information relating to that individual that can be found on the internet—information which potentially concerns a vast number of aspects

[2]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[3]Ibid., §93.

[4]Opinion Advocate General Jääskinen, 25-06-2013, C-131/12, ECLI:EU:C:2013:424 (*Google Spain SL, Google Inc./AEPD, G*), §131.

[5]Ibid., §33.

[6]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §35.

of his private life and which, without the search engine, could not have been interconnected or could have been only with great difficulty — and thereby to establish a more or less detailed profile of him".[7]

The case is still cause for discussion, with the discussants split between those that share Jääskinen's view and take search engines to be an objective intermediary, and those who agree with the CJEU and argue that search engines do something *additional* to what is already there. An additional issue here, is whether a search engine provider can claim the right to free speech with regard to the presented search results. Unfortunately, much of this debate quickly evolves to a dispute on a right to be forgotten *versus* the right to freedom of expression and information, giving limited attention to the question of *how* search engines affect the presence of our informational persona — let alone whether it is an issue that needs to be addressed. This leaves the evaluation of the problem, as well as the solution, often hanging in midair. I take the *Google Spain* case and the discussion surrounding it as a sign that an analysis of the manner in which search engines can affect the online informational persona is vital.

In this chapter, I will therefore examine how search engines affect the online informational persona. I will start my inquiry with discussing the role that search engines play in the Web itself. From there on, I again trace the impact of the technological mediation on the online assimilation of personal information in three directions: the production of the presented content, the construction of the presence of personal information, and the composition of the publics of the information. Lastly, I will conclude this chapter by assessing what challenges this mediation brings forth with regard to our informational persona.

For this analysis I will mainly focus on Google Search. The main reason for this choice is that Google Search is the biggest player on the search engine market in Europe.[8] Google's impact on the online information flow shows from the fact that its name has even become an English verb for using an online search engine (Diaz, 2008, p. 26). The second reason for choosing Google as main example is that Google Search was the search engine targeted in the *Google Spain* case.

6.2 Industrial gatekeeper of attention

While producing content on the Web is relatively easy, getting an audience can be challenging (van Couvering, 2008, p. 178). Due to the effortlessness and low entrance barriers to create content on the Web, the Web has grown into a massive collection of information. The user has to find her way around in this. Conveniently, the digital nature of online information allowed for the development of fast search systems that could index a massive amount of information and serve users on a nearly global scale through the internet. Search engines operators offer

[7] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §80.

[8] See Statcounter Globalstats, `http://gs.statcounter.com/search-engine-market-share/all/europe`, last accessed 21-04-2019.

users assistance with information retrieval by calling in machines to deal with the abundance of information and turn the mass of the Web into structured portion-sized chunks for consumption-on-request by users. Google Search sees it as its mission to "organize the world's information and make it universally accessible and useful".[9]

Serving a massive number of users with their personal information retrieval from a gigantic indexed database is not a minor task; a user would herself not be able to accomplish the retrieval manually to the same extent, let alone in the same time frame of mere seconds. By taking over and mechanising a big part of the information retrieval actions by agents on a large scale, search engines industrialise the retrieval of information in the tertiary memory (hence, search *engine* is a fitting name). By helping users to locate and access certain signifying objects in mere seconds, they heavily reduce the 'manual labour' required to retrieve information while accelerating the speed of the process.

The use of search engines is deeply integrated in the Web: many, if not all, graphical browsers offer 'search' in their navigation bar (see image 6.1). This feature consistently reminds the user of the quick assistance that can be provided by search engines to navigate the Web, while saving her the effort to navigate to a search engine's web page.

Figure 6.1: Search in the navigation bar of the Firefox Web browser

With this pivotal position, search engines heavily impact our access to online information. By mediating the retrieval of information, search engines steer our attention towards certain content (and away from other content) (DiMaggio *et al.*, 2001, p. 131). As such, search engines function as 'gatekeepers' (Nahon & Hemsley, 2013, p. 7). In this role, they play into the main value on the Web: attention (see section 4.5). They form a portal for attention and "work on the basis that they can turn any site into something only one click away from their search results, almost a subsidiary of themselves" (Fuller, 2003, p. 88). In this position, search engines centralise access to the Web (Fuller, 2003, p. 88).

However, search engines are gatekeepers on another level than the 'classic' mass media gatekeepers which generally consisted of small groups of professionals, like newspaper concerns, research institutes and governmental agencies; search engines are focused on the control of traffic instead of content (van Couvering, 2008, p. 177). Moreover, while traditional gatekeepers generally focused on a specific context or type of content, like news in a newspaper site and films in a film database, search engines transcend this 'classical' context selection by offering a vast array of various types of sources. A search engine offers search results referring to the public as well as the private, the old and the new, the local and the global, films, history, entertainment, fringe interests, and so forth.

[9] `https://www.google.nl/intl/en/about/`, last accessed January 2017.

Due to this overarching position, search engines are a gatekeeper of gatekeepers: they show us which newspapers to access for articles, which research institute to consult for the latest developments in computer engineering, etc. As such, the role of the traditional gatekeepers in setting the standard of what is considered to be valuable knowledge, has shifted to the search engine (Hinman, 2008, p. 68). This pivotal position gives search engines a significant power over the connection between audiences and content; audiences as well as content providers depend on this mediating technology to bring them together.[10] The more users and publishers rely on the mediation of a search engine to reach content, the more influential the search engine becomes (Pasquale, 2015, p. 14). Search engines therefore have a major impact on the online information flows and heavily affect the user traffic to web pages.

Furthermore, due to their overarching position, I argue that information retrieval by search engines has a certain decontextualised character: the user does not need to select a specific source context for her information retrieval. By gatekeeping at this overarching level and lifting the user's need to choose a specific contextual frame, the agency with regard to the contextual frame has shifted to the mechanisms of the search engine and "the very techniques of knowledge transmission have become the new gatekeepers of knowledge for the public in general" (Hinman, 2008, p. 69). The technological intentionality of search engines thus plays a pivotal role in the mediation between users and content.

However, as the technological intentionality flows forth from the design of the technology, and this in turn is shaped by the medium controller, it is again relevant to also have a look at the medium controller's intentions and interests. As the services of the search engine are offered for free to users (as we see on more places on the Web, see e.g., chapter 5), the revenue is made otherwise. The interests of the medium controller are often, as in the case of Google Search, (at least partially) commercial (Hargittai, 2000, p. 249). In the case of Google Search, the application is not only a search engine, but also an advertisement platform — or more precisely, an auction business; Google auctions advertisement space based on search query and profiling information of users (Zuboff, 2015, p. 97). Meanwhile, user behaviour is monitored and the resulting information is used to find ways to maximise revenue — a practice labelled as "surveillance capitalism" by Zuboff (2015). With such revenue models underlying the application, commercially driven search engines have an incentive to attract as much traffic as possible and compete for user attention with other applications and websites (Hargittai, 2000, p. 243). However, it is important to note that both website providers and users help stabilise the position and marketing strategies of search engines like Google Search, by using them, depending on them, and by deliberately employing strategies to optimise their display in the search results (Mager, 2012, p. 776). Much of this stabilisation of the position of search engines on the user side is likely to be attributed to ignorance and the acceptance of default settings (Mager, 2012, p. 777).

[10]The importance of this was highlighted by Marc Rotenberg in a presentation for the Tilburg Institute for Law, Technology, and Society on the 20th of January 2015 in The Hague.

6.3 Appropriation of content

In their search results, search engines offer users content that is made available online. By using the available online content in order to run its own service, the search engine commodifies content that is generated by others (Fuchs, 2012, p. 43). In this section, I will take a closer look at how this content is collected, and what this means for subjects.

Search engines gather the content from which they derive their search results, by making copies of Web content and storing these in their database (Pasquale, 2015, p. 7). However, the Web is too large and dynamic for search engines to fully index (cf. Gulli & Signorini, 2005). Thus despite the fact that search engines may aim to index the complete Web, they will only be able to index a (potentially significant) part of it. By indexing a part of the Web and a part not, a search engine necessarily engages in a selection of the sources that it indexes. As such, the assembling of the database is inscribed by design choices on inclusion and exclusion (Gillespie, 2014, p. 168). What is not indexed remains outside the scope of the search — and thereby of the searching user.

Search engines assemble their database with the use of 'Web crawlers'; these are bots that 'crawl' over the Web and copy the content and meta data of web pages (cf. Brin & Page, 2012). A Web crawler starts by visiting a set of URLs that it is given (the 'seeds'). From there on, the Web crawler is generally programmed to cover as much content as possible by following hyperlinks, while keeping a number of policies into account which are set by the medium controller (cf. Dhenakaran & Sambanthan, 2011). These policies program the crawler to prioritise copying certain types of sources over others. The content that is collected for a search engine's database, is thus heavily dependent on the design choices made with regard to the seeds and policies programmed in the crawler.[11] In the content collected by the crawlers, we thus see a significant expression of technological intentionality on the content of the search engine's database.

Website controllers also have control over whether their websites are indexed: websites may use NoIndex/NoArchive tags or a 'robots exclusion protocol' or tags in the HTML-document, also known as 'robots.txt', which prevent search engines from indexing the site.[12]. Being indexed by a search engine, is thus an opt out instead of an opt in. The default is that a site is indexed. Additionally, it is important to note that the affordances of digital information can easily frustrate attempts to prevent content from being indexed by adding robots.txt; since content can be easily copied and replicated elsewhere, there is a chance that it will still become indexable by search engines if it is republished at another location.

However, when people add personal content to the Web or participate in a online publication, they may not always take the indexing of this content by search engines into account. One of the issues resulting from the manner in which search engines use content published by others, is therefore the assumed consent of the

[11] The politics of Web crawlers is a research subject on its own and exceeds the scope of this study.

[12] See `http://www.robotstxt.org/`, last accessed 06-07-2019.

participants for this use (Tavani, 2016). This becomes increasingly problematic with content published by others, especially if, in turn, the website is controlled by a separate medium controller. Individuals may voluntarily participate in the creation of content on websites controlled by others, but this does not necessarily mean that they also agree with the indexing and display of that content in a search engine (Tavani, 2016). The interests of individuals participating in the generation of content and the website controllers can even strongly diverge. While the individuals may prefer to not have certain content indexed and displayed in a search engine, the website controller can prefer the opposite; given the pivotal role of search engines in connecting audiences with content, website controllers often welcome the mediation of search engines.[13] Many websites even have a financial incentive to attract as many users as possible. If a web page disappears from the search result, it is likely to experience a substantial decrease in incoming Web traffic (Grimmelmann, 2010b, p. 436). The result is that many website controllers gladly let search engines use their content for the search results without charge. The availability of online content by means of a search engine can thus entail a potential conflict of interests between an individual contributing to the content on a website and the site's controller.

The existence of this kind of conflict of interests becomes apparent in what I shall refer to as 'the BBC cases'. These cases refer to a list of 182 URLs of BBC articles that were initially displayed in Google Search in response to a name search.[14] In the period of July 2014 till May 2015, Google removed these 182 URLs as search result of a specific name query on request of the subject. BBC's managing editor McIntosh decided to voice the BBC's interest in the retention of these results in Google Search by publishing this list: "We are doing this primarily as a contribution to public policy. (...) We also think the integrity of the BBC's online archive is important and, although the pages concerned remain published on BBC Online, removal from Google searches makes parts of that archive harder to find".[15] The URLs are not fully removed from the database, but only from the results of search queries in which the name of certain individuals are used.[16].

When I examined these cases, I found that they relate to a wide range of topics and contexts (see Appendix A). Individuals not merely object to the display of search results in the case of crimes and misdemeanours, but also with regard to relatively unremarkable content. Examples are articles reporting on damage compensation given to car crash victims, a dispute about a lost dog, calls for help to locate missing persons, interviews with cancer patients and opinion polls on topics varying from games to politics. When assessing the cases, I found that of the 182 delisted URLs, at least 40 see to publications that came into being

[13] See e.g., the citation of McIntosh below.

[14] Neil McIntosh, "List of BBC web pages which have been removed from Google's search results", *BBC News*, 2014. `http://www.bbc.co.uk/blogs/internet/entries/1d765aa8-600b-4f32-b110-d02fbf7fd379`, last accessed 02-02-2019.

[15] Ibid.

[16] The BBC website states: "Update 29/06/15: Google has asked us to point out that links to the BBC articles below are only delisted from results for queries on certain names. They are not removed from the Google index entirely. We're happy to make that clear." Ibid.

due to the cooperation of the individual by for example giving an interview or participating in a discussion. In these cases, personal information concerning an individual would not have been encoded in the online information source if the individual had not agreed to cooperate. Despite agreeing to share their information at the time of publishing, these individuals wanted to have the hyperlinks to these BBC publications removed from the search results following a query on their name. Some publications were relatively new, so the problematic dimension of the trafficking of personal information by search engines seems to be broader than merely the retrieval of 'old' or 'outdated' information.

An exemplary set of removal requests involves the articles published in the series "BLLCKS — Check them. Don't lose them". As part of an awareness campaign for testicular cancer in October 2014, the BBC published interviews with six people on their personal experiences with this disease ("We spoke to five men - and a woman - about their relationship with their balls..."). In less than six months after the publication, five of the six interviewees requested the delisting of the interview in Google Search. Even though these individuals willingly gave interviews to the BBC, they objected to the display of those interviews by a search engine. Given the short time-lapse between the publications and the request for delisting, I conclude that it is not necessarily the signifying object in its original source that is experienced as a problem, but the manner in which search engines make the information present. I will examine how this presence is shaped in the next section.

6.4 Presence of personal information

In this section, I will the discuss the manner in which search engines affect the presence of personal information in the consecutive steps of the search process, starting with the user query, then the generation of the search results, and lastly the presentation of the results to the user.

6.4.1 Step 1: The query

A search starts with a user entering a keyword or set of symbols, the 'user search string'. When the user starts performing the query, we already see an expression of technological intentionality, albeit on a suggestive level. Many search engines offer two features that assist the user in formulating a search string that has a good chance of producing search results: autocomplete and autocorrect. The first is an *a priori* and the other an *a posteriori* 'suggestion' given by the search engine. These suggested formulations can influence the user's choice for the search query. I will briefly discuss both features and then discuss their impact on the presence of our personal information.

Starting with the *a posteriori* suggestion, autocorrect. After the search engine returns the search results, it occasionally offers the user suggestions for an alternate spelling of the query. The autocorrect feature can have various forms. Initially in

Google Search, it was displayed in the form of a hyperlinked question asking the user: "Did you mean [....]?". With this suggestion, the search engine invites the user to click on the link in order to perform a new search with the suggested search string. Currently, the query immediately changes to the more successful query, and offers the user to "Search instead for [query as the user spelled the search string]". The autocorrect feature helps users to perform similar queries with an alternate spelling or a corrected spelling mistake. For example, a search on "Paul Ricoer" changes the query to "Paul Ricoeur".

The autocorrect feature can be lucrative for search engines. As search engines tend to sell keywords to advertisers, they have an interest in the use of certain words and languages over others (Kaplan, 2014, p. 58). By offering similar search queries, autocorrect can transform a keyword without or with little value due to misspelling "into a potentially profitable economic resource" (Kaplan, 2014, p. 59). While autocorrect can certainly help out users, it is thus part of two information value schemes; one in which the keyword has a use value for the user, and one in which it has an exchange value for the search engine. Due to this difference, autocorrect could give rise to inaccurate expectations of the user with regard to the validity of her search string.

Because autocorrect steers the attention of the user, it can increase or reduce the presence of particular personal references, or even of full personae by steering the user to particular names. The main impact that it has, is that it steers the user towards public figures (e.g., when you search for "Karl Max", you will get results for "Karl Marx"). This can make it more difficult for users to retrieve information about people with an unusual name or spelling.

While autocorrect certainly has some impact on the presence of personal information, the impact of autocomplete is far more extensive. Autocomplete is a proactive feature that works during the encoding of the search string. It displays a list of possible search queries that start with the same letters or symbols that the user is entering into the query. The suggestions are shown in a drop down list as the user types in the search string. The autocomplete feature is displayed in action in figure 6.2.

Autocomplete suggests possible queries, generally based on a combination of a user's previous search history, language, popular searches by other users and trending topics.[17] In Google Search, personalised searches are prioritised in the autocomplete feature.[18] As such, the user's previous search history is the main shaping factor for her autocompletions. The suggestions that are offered to the user are calculated by algorithms.[19] If a particular autocompletion succeeds in grabbing the attention of the user, and she clicks on the completion, this can reinforce the completion's position as a popular suggestion and strengthen its position as a

[17] See `https://support.google.com/websearch/answer/106230?hl=en`, last accessed 7 March 2017.

[18] Danny Sullivan, "How Google Instant's Autocomplete Suggestions Work", *Search Engine Land*, 2011. `http://searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592`, last accessed 06-03-2017.

[19] See `https://support.google.com/websearch/answer/106230?hl=en`, last accessed 7 March 2017.

dominantly present annotation to the search string. However, the appearance of a certain search suggestion can also be short-lived. Some autocompletions spike at a certain point due to a specific event, but lose their presence as time passes by.[20] The feature often also takes a small margin of spelling error into account and corrects it.[21]

Users generally appreciate the autocomplete feature; it saves time because the user does not have to type in all the information, and the spelling suggestions are considered helpful (Ward *et al.*, 2012, p. 14). However, the most significant effect of autocomplete on the information retrieval process, is the fact that it points users to specific potential queries. Due to these suggestions, users experience autocomplete as "extra brainstorming, but from the computer" (Ward *et al.*, 2012, p. 12). Autocomplete can easily offer the user insight into informational relations of which she was unaware. It is therefore not just a completion, but a notification of the existence of certain informational relations established by other users. In this sense, autocompletes "offer a window into the collective Internet consciousness" (Baker & Potts, 2013, p. 201). It turns the user's attention to what is popular and calls into life a certain informational relation that can peak the user's curiosity to venture into the suggested direction. An example of this is shown in the query depicted in figure 6.2. It shows the suggestions offered when typing in the real name (crossed out in the picture) of someone who became publicly known as "Star Wars Kid".

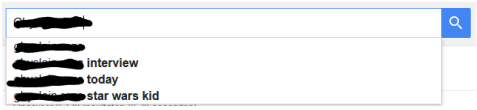


Figure 6.2: Star Wars Kid autocomplete in Google Search, the hidden text is the referent's name

An interesting autocomplete case to mention here, if only briefly, is the recent Dutch case of 'professor B'. This case revolves around a newspaper, the NRC (*Nieuwe Rotterdamse Courant*), that was prohibited by court to publish the full name of a professor who was accused of sexual misbehaviour.[22] While the newspaper was not allowed to publish the full name of the professor, some prevalent autocompletes following entries like "hoogleraar UvA" quickly reveal the man's name. This case is interesting because the case demonstrates how autocomplete can thwart court judgements.[23]

[20] Danny Sullivan, "How Google Instant's Autocomplete Suggestions Work", *Search Engine Land*, 2011. `http://searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592`, last accessed 06-03-2017.

[21] Ibid.

[22] Lineke Nieber, *Rechter: NRC mag naam van ex-hoogleraar niet publiceren*, NRC.nl, 2019. `https://www.nrc.nl/nieuws/2019/05/14/rechter-nrc-mag-naam-van-ex-hoogleraar-niet-publiceren-a3960173`, last accessed 14-06-2019.

[23] Additionally, the case gave rise to the 'Streisand effect', which I will discuss in the next chapter.

Moreover, the completions of autocomplete do not necessarily have an objective, nuanced or even truthful character: autocomplete is not limited to rightful informational relations, but *any* informational relation users have been interested in (Pasquale, 2015, p. 72). Autocomplete can therefore suggest illegitimate, wrong, harmful and discriminatory informational relations to users. It is prone to reproduce stereotypes, and can for example facilitate racism by suggesting that a certain relation exists (cf. Baker & Potts, 2013; Elers, 2014; Chander, 2016). Search engines try to address these issues by actively filtering autocomplete. For example, Google search aims to block suggestions that are related to hate or violence, to porn and adult content, to personal information like phone and social security numbers, to piracy and to suggestions that are legally ordered to be removed.[24] Also, in Google Search "[t]he autocomplete algorithm is designed to avoid completing a search for a person's name with terms that are offensive or disparaging. (...) This filter operates according to the same rules no matter who the person is"[25]. However, the filter does not prevent all illegitimate and discriminatory autocompletions.[26] Moreover, in many cases it will be a case of human — and more specifically court — judgement to decide whether a particular autocompletion is truthful or defamatory. Despite attempts to address these problems by blacklisting certain autocompletions, the rise of discriminatory and defaming autocompletions thus remains to be a problem. While defamatory and discriminatory autocompletions is a very interesting and socially relevant topic, I will leave the discussion here for what it is because this study has its focus on information that is not in itself illegitimate.

6.4.2 Step 2: Generation of the results

After receiving the user input for the query, the search engine generates search results. This takes place in a 'black box': the process preceding the search engine's output is hidden (cf. Fuller, 2003; König & Rasch, 2014; Pasquale, 2015). Hiding the manner in which the search results are generated, serves several purposes for the medium controller. By not giving insight into the manners of information processing, competitive search engine operators cannot copy or easily create a tweaked version of the search engine's processes. Moreover, it makes it difficult for other agents to game the search engine's ranking in their own interests (Pasquale, 2015, 64). Additionally, the lack of transparency is the result of what is commonly believed to be a 'user friendly' interface: hiding all the technicalities from view is said to make the use of a search engine more accessible for a bigger group of users (cf. Miconi, 2014; Campanelli, 2014).

[24]Danny Sullivan, "How Google Instant's Autocomplete Suggestions Work", *Search Engine Land*, 2011. `http://searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592`, last accessed 06-03-2017.

[25]Tamar Yehoshua, "Google Search Autocomplete", *Google*, 2016, `https://blog.google/products/search/google-search-autocomplete/`, last accessed 06-03-2017.

[26]See for example, *Man wins right to sue Google for defamation over image search results*, *The Guardian*, 2018. `https://www.theguardian.com/technology/2018/jun/13/man-wins-right-to-sue-google-for-defamation-over-image-search-results`, last accessed 14-06-2019.

Despite the fact that the generation of the search results takes place in a black box and the details remain unknown, some of the general elements have been disclosed by search engine operators themselves, as well as by researchers. I will discuss these general elements and their impact on the presence of personal information.

From the content that is indexed by the search engine, the search results are generated in response to the user search string. This entails, first of all, a technological interpretation of what the user is looking for. For the search engine, the user search string is a set of symbols devoid from any social or contextual connotation; the search engine is unable to treat the user search string in a contextualised manner from the user's perspective (Fuller, 2003, p. 71). In order to select search results from the search engine's database, algorithms are employed (see also section 5.5.2). These algorithms aim to return the 'relevant' search results. The designers of the algorithms play a crucial role in this: the relevance of information relies on often fluid norms that are open for interpretation "[a]s there is no independent metric for what *actually* are the most relevant search results for any given query" (Gillespie, 2014, p. 175). As the designers of the algorithms shape these according to their understanding of 'relevance', they imprint their normative views into the mechanisms of the search engine (Goldman, 2011, p. 107). The result is that the algorithms that evaluate the indexed content, represent "a particular *knowledge logic*, one built on specific presumptions about what knowledge is and how one should identify its most relevant components" (Gillespie, 2014, p. 168).

The most straightforward retrieval algorithm that is generally used to return relevant content based on a particular query, is the "Boolean approach". This is a true/false approach were "[l]inks to documents are returned only if they contain exactly the same words as your query" (Fuller, 2003, p. 83). A search engine is able to pinpoint the full or partial occurrence of the user search string in potentially massive signifying objects in its cache. However, the Boolean approach has its drawbacks: related sites not containing an exact (partial) match to the search string are not returned, because search engines are unable to deal with the variable understanding of words, like allegories, synonyms, metaphors and irony (Fuller, 2003, p. 84). The result is that search generally entails only a *literal* result retrieval. This is less of a problem when you are looking for individual names.

Next to an exact match to the search string, the algorithms generally also take the typeset of the search string in the original content into account. When indexing web pages for the database, search engines technologically evaluate and differentiate the content: "Each document is converted into a set of word occurrences called hits. The hits record the word, position in document, an approximation of font size, and capitalization" (Brin & Page, 2012, p. 7). Characteristic aspects of a text (e.g., title, bold font, bigger sizing) are recognised and interpreted by giving a certain 'type weight' to different fonts, sizes and function types of a specific (set of) word(s), as well as the proximity of words to each other (Brin & Page, 2012, p. 12). As such, salient elements in the original text are given more weight by the ranking algorithms.

Another kind of algorithm that can be used is Google's famous PageRank algorithm (cf. Page *et al.*, 1999). PageRank ranks a website based on the number of links to that website, as well as the estimated 'importance' of the website that does the linking (Page *et al.*, 1999). The more a website is linked to — especially by important others — the more 'authoritative' the website is taken to be and the higher ranked (Pasquale, 2015, p. 64). The goal of PageRank is to help users to "quickly make sense of the vast heterogeneity of the World Wide Web" (Page *et al.*, 1999, p.1). However, these mechanisms do not always work according to the humanly expectations of what 'relevant' search results are. For example, search engines do not differentiate between reasons why a website is linked to. Linking to a website can therefore have unwanted side effects: when users are linking a lot to a specific website as part of a critique, they can inadvertently give the website an authoritative status and turn it into a top search result (Pasquale, 2015, p. 73). In its core PageRank thus takes the attention value of a signifying object as factor to evaluate the importance of the content (Pasquinelli, 2009, p. 155). The consequence is that attention becomes the driving force behind the visibility and authority of web pages. This can lead to the promotion of commercial and popular websites over sources of information that are more detailed, noncommercial, and/or less easy to digest (Diaz, 2008, p. 13). Hinman therefore argues that search engines shift the assessment of the value of information from the traditional scientific and scholarly professionals to a technologically driven "digital version of the *vox populi*" (Hinman, 2008, p. 67). As such, PageRank backs a 'spectacular regime' in which "the value of a commodity is produced mainly by a condensation of attention and collective desire driven by mass media and advertisement" (Pasquinelli, 2009, p. 155).

Moreover, the mechanisms constituting PageRank tend to strengthen the position of the top websites by providing them with the highest visibility and thus the biggest chance to be clicked on and/or linked to. This leads to the 'Matthew effect' (a term coined by Merton (Merton *et al.*, 1968, p. 58)) in which 'the rich get richer and the poor poorer' (Origgi, 2012, p. 49). Origgi therefore labels PageRank as an 'aristocratic' network (Origgi, 2012, p. 48). Along the same lines, PageRank seems to prioritise older well-established websites: longer existing sites generally have more links to them than new pages which still need to build up their link 'reputation'. The result is that 'relevance' is often attributed to what is already popular and has a strong inclination to point towards English websites (Gillespie, 2014, p. 177), and websites from Western countries like the US and France.[27] The ranking mechanisms combined with the underlying commercial interests thus limit the types and sources of knowledge that are in the end presented to users, generally in favour of mainstream voices (Hess, 2008, p. 35-36).

Despite the prioritisation of older well-established websites, Google search does seem to have a certain contemporary focus in its generation of search results. Two

[27]Lily Kuo, "Almost all internet searches in Africa bring up only results from the US and France", *Quartz Africa*, 2017. `https://qz.com/1001555/biafra-the-threat-to-evict-igbos-from-northern-nigeria-is-being-swiftly-dealt-with-by-the-government/`, last accessed 10-06-2017.

bloggers discovered that websites originating from 2006 were not be returned as search result in Google Search, even if when using prefixes to specify the search and a search string that exactly matched the content.[28] One of the bloggers, Fioretti, found that this is different for other search engines like DuckDuckGo, which did display the result from 2006. The temporal scope of the displayed search results can thus differ fundamentally per search engine.

As the search engine selects and ranks content based on an algorithmic relevance evaluation, it expresses a strong intentionality in its information retrieval. By presenting 'relevant' search results, the search engine puts a 'relevance stamp' on the information flow. This is not unlike the effects of classic rhetorics where "[b]y the very fact of selecting certain elements and presenting them to the audience, their importance and pertinency to the discussion are implied" (Perelman & Olbrechts-Tyteca, 1969, 116). The search engine thus not only endows the content that it presents with a certain presence for the experience of the audience members, but it also imbues them with meaning.

6.4.3 Step 3: Presentation of the results

In response to the search query, the search engine offers the user an assemblage of search results that is produced by the search engine's algorithms. For the presentation of the results, the search engine frames the original content into its own context. In this subsection, I discuss how the presentation of the search results by the search engine transforms the presence of the references revealed by the original objects.

First of all, the original content is framed as a particular singular search result. The original object is generally turned into a signifying object consisting of a hyperlink and an image or zoomed-in fragment of the search string as it appears in the original object (see figure 6.3). With its focus on a literal (albeit partial or spread) occurrence of a search string, this zoom-in has a particular character: the search engine highlights the presence of the search string on a microlevel. Even a minor occurrence of the search string in a massive document can easily be displayed prominently. As such, the search engine distorts the context by providing a magnified view of the occurrence of the search string — not unlike a magnifying glass (as is often appropriately used as search pictograph). This magnification can give the user a distorted view of the actual positioning and relevance of the role of the searched for subject in the original signifying object. Search engines thus "are attention lenses; they bring the online world into focus. They can redirect, reveal, magnify, and distort. They have immense power to help and hide" (Grimmelmann, 2010b, p. 435). Moreover, as the zoomed-in partial representation allows the user to take notice of a part of the information about the search term on the search

[28] Tim Bray, "Google Memory Loss", *Blog*, 2018. `https://www.tbray.org/ongoing/When/201x/2018/01/15/Google-is-losing-its-memory`, last accessed 15-07-2019; and Marco Fioretti, "Indeed, it seems that Google IS forgetting the old Web", *Stop at Zona-M*, 2018. `http://stop.zona-m.net/2018/01/indeed-it-seems-that-google-is-forgetting-the-old-web/`, last accessed 15-6-2019.

result page, the user may not even need to consult the original content to find her answers (Pasquale, 2015, p. 59).

The singular search results are generally combined on the first page in blocked groups of images, advertisement, and text (see e.g., Google Search and Duckduck.go). In these various blocks, different sources can be leading. This division in blocks provides an overview of diverse angles and types of information in relation to a certain search string, allowing the user to see the occurrence of a search string in a variety of contexts at a glance. After the first page, we can stroll through a collection of the same type of search results, depending on whether we remain in the general tab, or select for example the images, shopping or video tab.

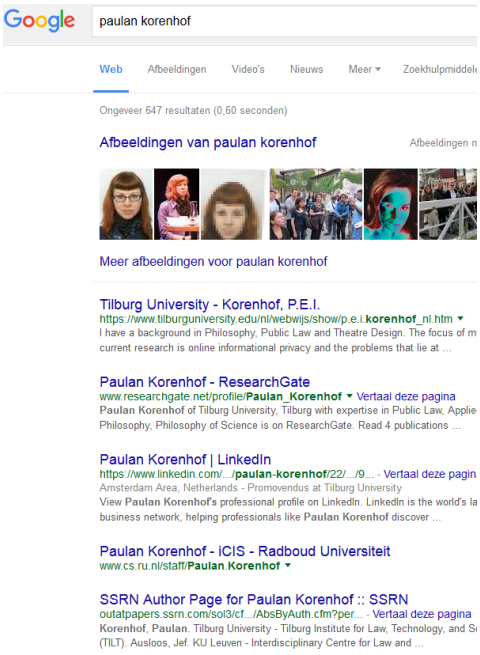


Figure 6.3: A Google search query with results

This ordered overview of search results offers the user the impression of comprehensiveness in the abundance of information of the Web (Vaidhyanathan, 2012, p. 59). The ranking presentation affects the presence of the references on a quantitative and qualitative level. The results on the first pages, and especially the top results, are the most visible to users and users tend to focus on these results (van Deursen & van Dijk, 2009). The top results are thus quickly taken in due to their prominent position on the page, but more importantly, the ranking of search results expresses an importance of information with regard to a specific search query. By attributing value to search results and presenting them to users, the search mechanisms are "producing and certifying knowledge" (Gillespie, 2014, p. 168). However, the production process of this 'knowledge' is not necessarily focused

on providing users with the most relevant knowledge. Instead, the composition and order reflect the technological intentionality of the search engine. In the search result overview, search engines present decontextualised 'popular fragments' under the banner of importance and objectivity (Gillespie, 2014, p. 179). Additionally, search engines like Google Search are designed to generate profit, and therefore also display sponsored results next to the unsponsored (often also referred to as 'organic'[29] results (Diaz, 2008, p. 20). By republishing a part of the signifying object in a new context, the ranked overview, and with different intentions than its original publisher, the search engine therefore expresses a certain technologically established perspective about the value and the meaning of the object for the searching user. Here, it is worthwhile to briefly take a sidestep to the essay *The Work of Art in the Age of Mechanical Reproduction* of Benjamin (2008). In this essay, Benjamin argues that when a camera captures an actor on film, while being "[g]uided by its operator, the camera comments on the performance continuously" (Benjamin, 2008, p. 17). As the camera reproduces the performance of the actor, it does this from its own perspective (i.e., the camera can provide close-ups, distant viewing, cut-outs, etc.) and thereby expresses a certain view on the performance — a 'commentary'. Along these lines, we can also say that a search engine 'comments' on online signifying objects by reproducing them according to the standards of the search engine's own framework: the search engine zooms-in, decontextualizes, and frames the content, thereby giving its own perspective on the meaning of the reproduced objects.

While the search engine thus expresses a relatively strong technological intentionality in the selection and presentation of the search results, its mechanisms underlying the assembly of the search result list are hidden from view for users. With a few exceptions, the user can only perceive the query and the following results.[30] At the same time, as black-boxed technologies, a search engine gives the user an illusion of control by inviting her to actively initiate the search (Sparrow *et al.*, 2005, p. 281). However, the user will never be sure to what degree she has control, or on what grounds the presented search results have been selected: what actually happens in 'the black box' remains inaccessible to her. Thus despite seeming user friendly, the opaque interface reduces the human agent to a blind operator and consumer of information that is filtered for her — subjecting her to imperceptible external rules and narrowing her choices (Lemmens, 2014; Rushkoff,

[29] With 'organic' search results, authors generally refer to those search results that are generated by the 'relevance calculating' algorithms without any interference to promote or demote certain results. However, I find the term 'organic' in this context somewhat problematic because it suggests a kind of natural process that is free from artificial interventions. For one, given the artificial design of the algorithm, can we truly speak of 'organic'? Secondly — and more importantly — given the black-boxed character of search engines, it is not possible to discern which search results are generated without manual interference. I have therefore chosen to use 'sponsored' and 'unsponsored' search results for those cases where I want to make a distinction between search results that receive their position in the ranking as a result of someone paying the search engine operator for this position, and other search results.

[30] One of the most notable exceptions was typically when Google Search has carried out a right to be forgotten request; in these cases it initially notified the user that search results have been removed due to European legislation.

2010). Because users have no knowledge of the selection processes in the search engine, they are oblivious to what may have been filtered out or never even made it to the search engine's index. As such, search engines like Google Search subject the user to information flows that are being defined — or rather calculated — for the user, while allowing little say to the original publishers of the content.

6.5 The individuated public

Web users depend to a great degree on search engines (Diaz, 2008, p. 13). In the massive web of online information, they have become the lifeline of users who are looking for specific resources. However, as tools for retrieval, search engines affect the effort and motivation underlying a search for online information with their easy accessibility and easy-to-operate characteristics. The effort needed to find online information with a search engine is by some perceived as "no work at all" (Downey, 2014, p. 141). The ease of the interface, and especially its implementation within browsers, invites the audience to use the search engine. As such, search engines give "anyone with a computer or a nearby public library access to resources that were once out of reach of all but the very few with unlimited funds and leisure time" (Pasquale, 2015, p. 60). The availability of search engines have turned the search for information from a mainly professional practice into a practice of the general public (Knight & Spink, 2008, p. 224). The use of search engines even seems to have become a social norm: "We are not simply enabled but also *expected* to use the search engine, in school, at home, and at social gatherings" (König & Rasch, 2014, p. 12). The search engine has been adopted by users into their homes, lives and implemented as part of their — often daily — routines. As such, search engines play a pivotal role as online gatekeepers that connect audiences with content. The consequence of their position is that search engines have "the power to ensure that certain public impressions become permanent, while others remain fleeting" (Pasquale, 2015, p. 14).

The connecting of audiences with content takes place in a 'two-way interface': the search engine provides the user with information, and the user (often unaware) provides the search engine with profile information that is being fed to the search engine in a feedback loop (Zimmer, 2008, p. 91). This allows for the personalisation of search results: the relevance of search results is fine-tuned for a particular user by combining the 'general relevance' with personal factors like the user's geographic location, language and/or previous search inquiries. This user profile information forms what Stalder & Mayer call the 'second index' (Stalder & Mayer, 2009). A Google engineer explains about the personalisation of search results:

It actually happens at every stage of the pipeline. When you start typing your query, if you're signed in, the autocompletions will prefer queries that you've typed in before. If you're in a given metro area, we will prefer queries that make sense to you in that metro area.The second level it happens at is, when we process your query, we also take into account your Web history and so on in order to guess

at your intent. During ranking, the process of actually looking at the documents, we also take into account personal signals that make sense for you, and when we search for your personal content in Search, plus Your World, we take into account your personal signals over there. Finally, when we have the full set of results assembled, we then customize them for you.[31]

This kind of personalised filtering can enclose the user in her own little information universe, where the information that does not fit her profile never reaches her attention; the user is then enveloped in what Pariser calls the 'filter bubble' (Pariser, 2011). Due to the profiling "the user becomes prescribed in their experiences" (Hess, 2008, p. 35). With every click of confirmation, she is more strongly enclosed in a particular informational profile that, literally, forms the protention for her future interactions with the online tertiary memory.

As search engines connect audiences to content based on a personalisation of the user combined with the ranking algorithms, they both expand and limit the spread of information to different publics: they can connect audiences with content that these audiences would otherwise never encounter, but at the same time they differentiate between audiences and make a connection between audiences and content based on the profile information of the user. As such, search engines segregate audiences — but according to the logic embedded in their algorithms. This can be a different segregation of audiences than the original publishers of online content may had in mind. For example, imagine a blogger who has written an in-depth article about depression aimed at a worldwide audience and posted this on her website with a .nl extension. For the article, she interviewed five people from the Netherlands. Because of the characteristics of her website, there is a fair chance that mainly only an audience with a Dutch IP-address is directed to her blog. For the blogger this is an unwanted limitation of her audiences. Moreover, because the blogger is not an authoritative source, it is doubtful her blog will end up high on the ranking for people searching only for 'depression'. However, because the interviewees are people with a limited online presence, her blog is likely to end up high in the ranking in return to queries with the interviewees' names. For this blogger, this entails an audience segregation based on the wrong relation between audience and content. As such, the audience composition by a search engine can establish a new and different audience segregation with regard to online content than the original publisher may have in mind.

Search engines thus heavily affect the information retrieval of their users, as well as the traffic to content on websites. However, it is important to note that the impact of the search engines on our information retrieval, ties in with the manner in which users tend to use them: it is the user who heavily uses search engines, tends to focus on only the top search results, hardly look past the first search result page, and spend little or no attention on the source of the search result, nor the date of the information (van Deursen & van Dijk, 2009).

[31] Jon Mitchell, "How Google Search Really Works", *ReadWrite*, 2012. `https://readwrite.com/2012/02/29/interview_changing_engines_mid-flight_qa_with_goog/`, last accessed 03-03-2019.

6.6 Complications of the presented persona

In this final section of the chapter, I shall discuss how the mediation of information retrieval by search engines can raise complications for our informational persona.

Search engines play a pivotal role in the access of online information. Being part of the Web, search engines have similar access requirements and affordances as basic websites with regard to the information they present (see chapter 4). However, because search engines offer users assistance with the *retrieval* of information on the Web, they take in a fundamentally different position than basic websites and social media. As a third party processing the content of other websites and presenting it to users, search engines have industrialised the retrieval of online information and established a particular technologically mediated retrieval praxis. With this particular position, the search engine has a specialised power position as gatekeeper.

Despite the fact that users play a vital role in the retrieval praxis of search engines by being the ones who trigger the machine with a particular query, I argue that the user is not the root of the problems. We may have valid reasons to search for information about someone. To give an example in the case of a name query, you may want to look up an article written by a speaker you just saw at a conference. In these cases, you are interested in certain aspects of the person. For the human user to 'search' is a specific act with an intentional directionality towards a particular object of interest. However, this is where the mechanisms of the search engine kick in and impress their own intentionality on our information retrieval action: when the search engine performs a search on our request, it translates this act into its own technological modus.

First of all, due to its technological modus, search engines do not guarantee — or even check — whether the content they present is correct. On this level, search engines share a part of the issues they can raise with the Web in general: as the search engine automatically collects all sorts of content on the Web, it can easily index content that is inaccurate, (intentionally or unintentionally) false, and in turn present faulty information about an individual to users. This can even be an abundance of false information, when the individual was for instance the victim of a smear campaign. What complicates this even more is the 'truth effect' that can be caused by the repetition of information (Sparrow *et al.*, 2005, p. 281): if the same (kind of) information is displayed as search result on successive searches, users are more inclined to perceive the signifying object as 'truthful'. The retention of a user's profile and search history could increase the chance of repetitions. The role of the search engine is clearly problematic if the search engine composes a perspective on the informational persona that contains faulty information. However, the problematic impact of search engines runs deeper than the possible display of faulty content.

The moment the user starts entering her search string, the search engine already becomes actively involved in the act of searching by offering the user suggestions by means of autocomplete. By steering the user attention with autocomplete, certain personal information, mainly the popular and the recent trends, can be

highlighted and brought under the attention of users. Autocomplete affects the user's perception of the persona by establishing associations and sometimes even led to unjustified connections between an individual's name and other terms. An example that shows the problem of this, was the autocompletion of the name of a former German First Lady. She was the victim of a false rumour that she had worked as a prostitute. As this spiked the general public's interest, users tried to search online for information with the help of Google Search. The result of this public search was that when typing in the First Lady's name, it was autocompleted with terms like 'prostitute', and 'escort'.[32] Also, autocomplete can reveal a referent's identity, as it did in the case of professor B discussed in section 6.4.1. In such cases, the search engines' autocomplete function facilitates a reverse name query.

Once the search engine received the user input (whether it be a new search string or a click on an autocompletion), it recontextualises the search act and its directionality into its own frame of reference. The meaning or context that the user may have had in mind is replaced by the mechanisms of the search engine that combine the user profile with the search engine's ranking and selection processes. By interpreting the search string as a flat set of symbols and matching it to items in the index, the search engine focuses on the retrieval of literal matches to the search string, potentially even at a microlevel. It is here that we find one of the biggest implications for the informational persona: the microlevel retrieval allows users to locate signifying objects in which a personal name occurs even in the smallest details. For example, search engines can single out the opinion of a particular person in a big and/or remote online discussion on for example animal testing or a reality-tv show.[33] Moreover, the microlevel retrieval of a name can cross language, cultural and even script differences. An example of this, is a case where an article in Cambodian local media reached a US audience by means of a name search. When a US citizen died on holiday, the local Cambodian media published his name and autopsy picture online. Despite the differences in script (except for the name, the full article was in Cambodian script), language and a different national territory, the Cambodian article and autopsy picture were returned as results following a search on the name of the deceased.[34][35]

As search engines zoom-in and connect audiences with content based on a search string in this manner, they actively reconfigure the original content into a new perspective by reshaping the context and redefining the conditions of retrieval

[32] See Stefan Niggemeier, "Autocompleting [...]: Can a Google Function Be Libelous?", translated by Paul Cohen, *Spiegel Online*, 2012. `http://www.spiegel.de/international/zeitgeist/google-autocomplete-former-german-first-lady-defamation-case-a-856820-druck.html`, last accessed 02-02-2019.

[33] These are examples of opinion pages for which Google Search received a request for the delisting of the search result. See the BBC cases, Appendix A.

[34] Joe Nocera, "Try a Little Common Sense: Some Material Ought to Be Delinked by Google", *New York Times*, 2014. `http://www.nytimes.com/2014/06/14/opinion/joe-nocera-some-material-ought-to-be-delinked-by-google.html?emc=edit_tnt_20140613&nlid=28836431&tntemail0=y&_r=0`, last accessed 2017-03-07.

[35] While the GDPR does not see to the protection of information relating to the deceased, I find this an exemplary case to explain the global cross-culture mechanisms of search engines.

of the original signifying object. In this, they can easily frame the content in ways that diverge from the intentions of the original publisher. Search engines can construct new audiences for online publications, which differ from the original intended audiences of the publisher (or at least the participating individual, see section 6.3). While agents can be open about particularly personal information to the general public, this openness is generally context related, e.g., in the context of a discussion or research on people's experiences with health care, discrimination, relationships, school or the like. The audience that such content attracts is commonly the audience interested in the *topic* that is being discussed, not in the particular individual. Search engines reconfigure the audience-access relation by potentially flipping this interest around, especially when performing a name search, thereby reversing the incentive needed to access the information as well as the focus of the original content — and attaching the topics to the individual's informational persona as predicates. In particular, the focus reversal afforded by a name search de-anonymises one or more particular informational persona(e) from the mass of information and deforms the context in which the information was originally shared by highlighting its individual presence. The zoomed-in focus of search results can thus easily decontextualise information, or turn the marginal or what is merely a side issue, into a headline. As such, the search engine as technological retrieval mechanism functions as a spotlight: by switching the spotlight from the main character to someone in the chorus in the background, the search engine can direct the attention of the audiences to the secondary and present it as a leading element.

The result of the retrieval praxis of the search engine is that even if the original content in its original location is unproblematic, and possibly even authored by the referent herself, the appropriation of this content by the search engine may give rise to problems, as the above-mentioned example of the testicular cancer interviewees demonstrates. What adds to this, is that search engines do this without having received explicit consent of the website controller that hosts the signifying object, the publisher of the object (if that is someone else than the website controller), or consent of the referent to whom the signifying object refers. By commodifying and decontextualising the content and placing it out of control of the author, the search engine appropriates the content and separates it from the author. This process alienates the author from 'her' content in the search engine. The only way in which the author can impact the manner in which the content is displayed as search result, is by adjusting the content on her own page (in which case the cycle repeats itself because the author still does not have direct control over the display of her content) or by the use of robots.txt. Search engines thus create a perspective on the informational persona that runs relatively loose from human steering by any of those involved in the publication of the original online content. I say 'relatively', because on the one hand, the algorithms are created and tweaked by human designers and engineers, and on the other hand, there is a lively market for 'search engine optimisation' (SEO) in which users and specified companies aim to influence search results by playing into the ranking algorithms in order to uplist or downlist specific URLs.

Furthermore, when search engines reshape original content into a new signifying object with a particular perspective, the search result, they do more than produce a single result: they produce a collection of search results. In case of a name search, search engines potentially detail vast amounts of information about the individual in one overview.[36] As the zoomed-in objects in various formats and originating from diverse sources are collected from the database and combined in a search result list, they can provide the user with a broad overview of personal references ready at hand for inspection in the here and now — potentially covering anything from the old to the new and from the public to the private. The profiling of users increases the success of a spot-on search focused on a particular individual: by aligning meta information, the results are likely to be focused on one or a few main informational personae that match with the user's profile. We can see this confirmed by the fact that when users use their own name as a query, the top-ranked results generally refer to them, and not to a person with the same name (Pasquale, 2015, p. 78).

Additionally, the collection of the search results affects the references that it contains. When objects are displayed in proximity to each other, relations between these objects are established (Mayer, 2009, p. 68). With this, search engines add new value to the results; "[w]hat search uncovers is not just keywords but also the inherent value of connection" (Kelly, 2007, p. 90). This combination of objects affects their mutual interpretation by turning them into each others' context. Moreover, this combination is more than just a collection: it is a ranked selection. In this, the search engines function as an authoritative voice in an external position to the objects and 'comments' on the signifying objects that it indexed. With its ranking based on a evaluation of the attention value of the content and its source, the search engine implies the importance of certain content for the search string. The higher ranked particular references are, the more prominently present they become as predicates for the query.

By making all these decisions on the level of zooming in, selecting, ranking, and turning the referent into the topic, search engines compose a particular perspective on the informational persona for a searching audience. With this, the search engine makes a certain claim about the referent's identity. This claim entails a perspective on a referent's informational persona that is broader than any single object reflects. However, it does so at the cost of depth and context: only snippets of objects are presented, zoomed-in and centred around a particular reference, with those with the highest attention value made the most prominent. The search engine thereby presents the persona as an exhibition. In this persona exhibition, the search engine reconfigures the existing informational relations of the original content (time, social connections, context) into the technology's relevance calculations. The social- and technological effects that play a role in the original sources, like

[36] The success of a search for signifying objects concerning a specific individual is thus anchored in the user search string. In case of a name change (for instance, due to marriage), the retrieval of signifying objects concerning that particular individual is impeded because the common denominator is missing; signifying objects not matching the name used as search string are likely not returned as search results.

the dropping of content to the bottom of a page with the passing of time and the expectations of the original authors with regard to the audience, may be bypassed, crossed or even nullified. Meanwhile, the zoomed-in presentation of the search result poses a challenge to the historical and socially contextual understanding of the content by the searching user. The displayed fragments become part of her *informational present*, while temporal context can easily be lost to her (see section 6.5).[37] Search engines can thus easily focus the attention of the audience to minor, private, or outdated aspects of an informational persona, and blow their meaning out of proportions. Especially in the case of individuals with a limited online informational presence, a specific signifying object can become a salient aspect of their informational persona due to the authoritative status of a particular source, like the content of popular media or an online newspaper archive. As such, search engines can "set a *spectacular value* for anything and anybody" [emphasis in original] (Pasquinelli, 2009, p. 159). This is also what likely happened in the *Google Spain* case: relatively old information made it to the top search results as a result of the authority of the source (see section 6.4.3).

The consequence of this all is that search engines can present a perspective on the referent's informational persona that diverges (sometimes even fundamentally) from what the referent feels is relevant for her identity. To be problematic, this perspective does not have to see to information that is noteworthy or special in one way or the other — simply being outdated can be sufficient to put forward a portrayal that can be experienced as problematic. As a referent explains: "When I was 20 years old, I made a website for a college course about building a digital identity. Today, it makes me cringe—largely because the site has become such a stubbornly resilient piece of my digital identity. (...) The site still appears—in all its lilac and teal glory—on the first page of search results whenever anyone Googles my name. A family video the whole world can see"[38]. Meanwhile, the impact of this representation can be severe: it can impact the referent's identity in the eyes of others, while in turn their reactions and responses towards the referent reflexively shape her self-understanding. For example, in the case of the referent quoted above, her amateur site could easily undermine her now professional career as a journalist by representing her as an amateur. Moreover, in the case of a vanity search (a referent using a search engine to find information about herself), the search engine itself takes in the place of a reacting other, by showing the referent its ranked view on her, thereby possible triggering memories and reflexively affecting her self-understanding. As the mediating technology presents who the referent is to others and the referent alike, it alienates the referent from her own history; her informational history is appropriated, while she herself has no control with regard to how she is represented by the search engine. Search engines can thus

[37] Yet, it is important to mention here that the original object may itself blur its relative age. The meta data of signifying objects may mark them as being created on the upload date, while in fact they may be much older. This would especially be a potential issue with regard to old analogue archives that are now scanned in and uploaded.

[38] Kaitlin Mulhere, "An Embarrassing Website I made in College Has Followed Me for a Decade. Here's How I Finally Erased It From My Google Search Results", *Money*, 2018. `http://money.com/money/5441177/manage-google-results-online-reputation/`, last accessed 25-04-2019.

make it difficult for individuals to successfully distance themselves from marginal peculiarities or past views and actions in the eyes of those who use the search engine. They thereby can undermine an agent's self-determination. While some users may be able to use the perspective that the search engine offers to their advantage by allowing them to catch a glimpse of how others perceive her in digital form and use this to (re)construct her identity (which, after all, is shaped through the eyes of others), this beneficial use is limited to a relatively passive check and does not change the manner in which the search engine presents the referent. The manner in which search engines convey information to others thus in general challenges the autonomy of agents with regard to how and when to convey certain personal information to others, as well as the possibility to provide additional context in order to address misunderstandings or wrong interpretations. This may potentially hamper second chances and the personal development. The extent to which search engines can confer a particular view of a subject to audiences can even be the cause of great distress because subjects may feel that they are unable to move beyond a certain view that the search engine presents of them (Ronson, 2016, p. 211).

From the applications that mediate personal information discussed so far, search engines express the strongest technological intentionality in their presentation of the informational persona. The search engine tells users what is important about the referent. Even more, as the use of search engines seems to have become a social norm, the information they present is likely to have a strong and relatively dominant presence. The consequence is that the perspective of the search engine becomes a particularly authoritative voice in the portrayal of the online informational persona.

Chapter 7

Going viral

Contents

7.1 Introduction

JUST REMEMBER! THE TECHNOVIKING DOESN'T DANCE TO THE MUSIC! THE MUSIC DANCES TO THE TECHNOVIKING![1]

Many Web users will have come across the phenomenon 'Technoviking', or a reference thereto, somewhere during their onli*f*e. Technoviking is one of the famous concepts brought to life by Web culture. Originally, Technoviking consists of video footage from a streetrave in Germany — the so-called 'Fuck Parade'. In the video a tall muscular man with a beard, long hair in a braid and bare chest dances to technomusic. In the beginning of the video, he scolds a man for harshly bumping into a woman by pointing an outstretched arm and finger upwards. The appearance of the imposing man, especially his scolding posture, became an iconic image on the Web.

The video was initially shot in 2000 and uploaded on the Web under the name 'Kneecam No.1' as part of an art project.[2] Here the video stayed relatively 'dormant' until about 2007. When in 2007 the video was picked up on a forum and dubbed as 'Technoviking', it rapidly gained widespread popularity and was massively shared online.[3] The image of the 'Technoviking' was used on a large scale for the creation of online art, remixes, jokes, re-enactments and parodies of the video. Next to that, the image of the man was used as a print for T-shirts, wall stickers, action figures and art projects — of which the most striking is a massive inflatable Technoviking head which fills with a little bit of air every time someone tweets '#technoviking'.[4]

Because Technoviking became popular trough a quick spread from user to user, we say that it has gone *viral*. With 'viral' in this chapter I am thus not referring to something relating to or caused by a computer virus, but a phenomenon which consists of information being "quickly and widely spread or popularized especially by person-to-person electronic communication".[5]

The virality of 'Technoviking' did not sit well with the referent, the original individual who was dancing on the street. He started a court case against the initial publisher of the video to have the material removed.[6] According to the referent, the video had severe negative consequences for his professional life.[7]

[1] `https://youtu.be/FwsntHcWiy4`, last accessed 20-09-2017.

[2] Matthias Fritsch, "Technoviking Archiv", 2000-2019. `http://www.technoviking.tv/subrealic.net/works/installation/technoviking-archiv/archive.html`, last accessed 20-09-2017.

[3] Ibid.

[4] Emma Hutchins, "Technoviking Meme Resurrected as Giant Tweet-Powered Head", *Mashable*, 2012. `https://mashable.com/2012/09/05/technoviking-tweet-powered-head/\#jZxiDny0aZqZ`, last accessed 20-03-2018.

[5] https://www.merriam-webster.com/dictionary/viral, last accessed 20-09-2017.

[6] Landgericht Berlin, 30-05-2013, Nr. 27 O 632/12.

[7] Ibid.

Technoviking is not a stand alone-case of an individual experiencing the consequences of virality. Many individuals have been the subject of a viral information flow. Some famous examples of viral cases are the 'Star Wars Kid', the 'Dog Poop Girl', and the 'Overly-Attached Girlfiend'. All these cases have in common that their content is relatively harmless, i.e., no explicit content, sex or violence, and the content can not be considered libellous — which make these cases and their implications especially interesting for this study. I will briefly discuss these three cases to give the reader some idea of the variety of viral cases that revolve around a particular person. They will serve as backdrop for this chapter.

Providing a perverted pleasure for Web users, the teenager who became known as the 'Star Wars Kid' was "a miserable and unwilling star of what media activists and analysts like to call 'user-generated culture'"(Vaidhyanathan, 2008). Star Wars Kid is a teenager who filmed himself in 2002 swinging a golf ball retriever while pretending to be a Star Wars jedi fighting with a light sabre. The teenager made the cassette tape on his high school for his private use. However, he forgot to take the cassette home with him and eventually classmates found the footage and uploaded it on a peer-to-peer network in April 2003 (Solove, 2007, p. 45). Here the video was picked up by a user who edited the video by replacing the golf club with a light sabre and adding sounds (Solove, 2007, p. 45-46). This video was then picked up and published by a blogger, who published the edited as well as the original version of the video and named the video 'Star Wars Kid'.[8] From that point on, the video started circulating across websites and gave rise to art, remixes, re-enactments and parodies. The video was not only massively viewed and shared, but also often accompanied with negative comments about the teenager's appearance (Solove, 2007, p. 46). Due to teasing and harassment on his school as well as online as a result of the Star Wars Kid video, the teenager suffered deep psychological distress, left high school and came under psychiatric care. When interviewed in 2010 the 'Star Wars Kid' recalls the virality period of the video as "a very dark period"[9].

An example of another type of viral content that is worth mentioning, is the 'Dog Poop Girl'. This video originated in South Korea, which has a relatively strong cultural focus on shame (You, 1997; Lee, 1999). When a woman's dog pooped in a subway, another passenger asked her to clean it up. The woman refused. This interaction was filmed, uploaded online, and went viral quickly after the initial uploading (Solove, 2007, p. 1). The content was spread to encourage the general public to condemn the woman's behaviour (Dennis, 2008, p. 351). As a result, the woman became known worldwide as the 'Dog Poop Girl', and was publicly shamed on a global scale (Vaidhyanathan, 2008).

A completely different kind of virality, the last one in this set of examples, is 'The Overly-Attached Girlfiend'. The 'Overly Attached Girlfriend' is a video made by the referent herself, in which she performs a parody on a pop-song from

[8] Andy Baio, "Star Wars Kid", *Waxy*, 2003. `https://waxy.org/2003/04/star_wars_kid/`, last accessed 30-03-2019.

[9] "10 years later, 'Star Wars Kid' speaks out", *Maclean's*, 2013. `https://www.macleans.ca/news/canada/10-years-later-the-star-wars-kid-speaks-out/`, last accessed 26-03-2019.

a stalker perspective. When she uploaded the video it went viral and gave rise to a stream of remixed signifying objects. The referent benefited from the virality by using her viral status for business purposes. In an interview she stated about the virality of her video: "It's definitely weird...but it's fun. I like it a lot"[10].

With these cases in the background, I will examine in this chapter how virality affects the online informational persona. Because virality is a phenomenon that theoretically can occur to any of the signifying objects discussed in the previous chapters, this chapter builds on their findings and therefore has a somewhat different character. I will start my inquiry by first taking a closer look at the phenomenon of virality itself: what is a viral outbreak? After that, I will look into the mechanisms of a viral outbreak and how it affects the presence of the reference that is at the centre of this outbreak. For this, I will first discuss the role and impact of the audience in virality because they play a key role in order for content to go viral. Next, I will discuss how virality impacts the presence of a particular reference. Lastly, I will conclude this chapter by discussing what complications a viral outbreak raises for our informational persona.

7.2 Online virality

Viral content is not just content that managed to attract collective attention, but content that also provoked users to spread it further in one form or the other. Nahon & Hemsley describe 'virality' as "a social information flow process where many people simultaneously forward a specific information item, over a short period of time, within their social networks, and where the message spreads beyond their own [social] networks to different, often distant networks, resulting in a sharp acceleration in the number of people who are exposed to the message" (Nahon & Hemsley, 2013, p. 16). According to Shifman, the key characteristics of viral information are a "(1) person-to-person mode of diffusion; (2) great speed (...) and (3) broad reach" (Shifman, 2013, p. 55). Despite its strong social connotation, the viral phenomenon also has a technological side: the conductive affordances of the medium affect the when, what, how and why of information sharing.

Given the character of the online environment, the reach of a viral event can be split in two elements: "(i) *reach by numbers*, the reach in terms of the number of people exposed to a content; (ii) *reach by networks*, the reach in terms of the distance the information travels by bridging multiple networks" [emphasis in original](Nahon & Hemsley, 2013, p. 29). The specifics of what entails a 'viral' distribution can depend on many factors like the total number of users who come in contact with the information, and the speed and spread of the distribution (Nahon & Hemsley, 2013, p. 16). While there seems to be no consensual definition of when exactly we can consider information to have gone viral, the three key

[10] Alyson Shontell, "The Overly-Attached Girlfriend Explains What It's Like Being A Wildly Popular Internet Meme", *Business Insider*, 2013. `https://www.businessinsider.com/the-overly-attached-girlfriend-explains-what-internet-stardom-is-like-2013-3?international=true&r=US&IR=T`, last accessed 19-03-2019.

characteristics listed by Shifman (above) seem quite useful as working definition of virality.[11] Additionally, for the purposes of this study, it is not vital whether or not a specific case can be considered as really having gone viral or not. Instead, the mechanisms and their potential consequences can tell us much about the problems that potentially result from the spread of a specific reference, despite whether the scale is sufficiently large to be labelled as viral.

In the following sections, I will dissect virality to get a better understanding of the phenomenon and the role played by its components. For this, I will start by taking a closer look at the driving force of virality: online publics.

7.3 The republishing audience

Virality can in theory follow as a result of the online accessibility of any of the signifying objects as discussed in the previous chapters: the only thing that is needed, is that audience members pick the content up and start forwarding it excessively. A vital element of virality is a public that comes into action.

In order to go viral, a signifying object first of all needs an audience: enough people need to (want to) see the object, before they will even consider to forward it to others. Obviously, the more strongly that a particular reference is already present, the bigger the chance that it catches the attention of an audience. Given the abundance of online information and the limited attention capacity of users (see section 4.5), there are certain characteristics of content that increase the likelihood of particular content grabbing the attention of a wider audience: viral content is generally simple on all levels (Shifman, 2013, p. 81), and easy and quick to digest (West, 2011, p. 83). Pictorial objects are therefore more likely to go viral than text objects. Also, pictorial objects are generally understandable by a wider audience (see also section 4.5). Moreover, the content is often focused on one narrative: e.g., 'man dances with very characteristic moves in a streetrave', or 'a boy plays with a golf ball retriever as if he were a jedi'. The packaging tends to be clear, straightforward, accompanied by snappy titles (three words or less) and in the case of videos, the object commonly has a relatively short runtime and a high degree of repetition (Shifman, 2013; West, 2011).

This corresponds to the view of Varis and Blommaert, who argue that the core of virality does not lie in the meaning of the content, but instead in its effect (Varis & Blommaert, 2015). It is part of a phatic form of interaction: the sharing of content serves more as a social action than as a sharing of information (Varis & Blommaert, 2015, p. 41). This social action can have different forms.

To start with, a viral forwarding can be the result of a wish for social interaction or bonding. Especially comical content can invite social bonding as "[s]haring humor signals similarity — and similarity breeds closeness (...) [l]aughing together is a sign of belonging" (Kuipers, 2009, 219). A related reason for users to transmit

[11] There is a similar debate on when something can be called 'Big Data'. This concept is also hard to define, but seem to have stabilised on the view that the key factors for something to be regarded as 'Big Data' are the volume, variety, and velocity of the data.

content, is to increase their social status and present themselves to others as having certain views, interest or a sense of humour (Teixeira, 2012). The result is that content that is experienced as 'pleasant' by the viewers tends to be more prone to virality than 'unpleasant' content (Eckler & Bolls, 2011).

Virality can also be the result of a directed social action: a viral spread can be intentionally provoked by a user or a group of users. Users can try to achieve an outbreak by strategically inserting specific signifying objects in the information flow and/or by attempting to affect the object's ranking on certain websites (Burgess, 2008, p. 104). They can have various reasons for wanting to provoke an outbreak: they can do it for fun, commercial interests, punishment[12], or to counter censorship. Dog Poop girl is an example of a viral public punishment. In this case, the public forwarded the content in order to publicly shame the referent.

The intentional push of a viral spread to counter censorship is of particular interest in the light of this study, because it is a reaction to an attempt to remove certain signifying objects — like a removal following an art. 17 GDPR request — and results in the opposite effect. The (attempted) censoring of content can thus attract the interest of users and provoke a viral outbreak (Nabi, 2014). A famous example of this is the case of Barbra Streisand who attempted to have photographs of her house taken offline. Her attempts to enforce the removal of the photographs by means of a lawsuit, spiked the media's interest as well as that of individuals who criticised Streisand's actions. This backfired and resulted in a broad coverage on the issue as well as a massive distribution of the photographs on the Web. Following these events, the phenomenon of causing a viral outbreak by trying to suppress or censor that very content, has been dubbed the *Streisand effect* (Nabi, 2014).

By triggering a social and relatively unified informational sharing-and-response wave, a viral outbreak establishes a kind of "temporally bound, self-organized, interest network in which membership is based on an interest in the information content or in belonging to the interest network of others" (Nahon & Hemsley, 2013, p. 34). Those unaware of the viral content are not part of the interest network and cannot 'join the conversation'. As such, viral content is not merely widely shared, but as viral content, it itself also attracts audiences: people want to join in on the conversation and see what the fuss is about (Nahon & Hemsley, 2013, p. 78).

However, the social character of the event is only one part of the motivation that drives an audience into a viral event: generally, the content itself sparks an emotion in the audience that trigger users to forward to content. The emotions that particular content can evoke matter because not all emotions equally trigger a user to forward content (Berger & Milkman, 2012). For example, content that evokes sadness is less likely to go viral because sadness tends to be a deactivating emotion, while content that evokes awe (positive) or anger (negative) is more likely to go viral because these are emotions that tend to arouse or activate a user

[12]From what I have encountered, viral punishment generally seems to focus on misdemeanours, or animal or child abuse. I did not come across cases where acts like murder were virally punished. Maybe this is because the public expects the government to execute the punishment in these cases. However, so far I have not found conclusive evidence for this.

(Berger & Milkman, 2013, p. 21). Research has shown that commonly the main emotional reaction that activates the forwarding of content is surprise (Dobele *et al.*, 2007; Dafonte-Gomez, 2015; West, 2011; Teixeira, 2012). The result of the dependence on a driving force like surprise and other activating emotions, is that content that goes viral often reflects a certain irony, portrays common people accomplishing impressive tasks, or do something that is contrary to the stereotypical first impressions that they make (see e.g., West, 2011). This is further underlined by the fact that viral content often features non-famous individuals (Shifman, 2013, p. 74), and is made by amateurs (Jiang *et al.*, 2014).

The forwarding audience affects the content: in their (re)encoding of the reference, they generally name and frame it by means of comments and the like that emphasise a particular social response to the content (e.g., annotating the content with a laughing or angry emoticon). As such, the forwarding of viral content is likely to place it in a certain 'social wrapping'. This social wrapping is generally unambiguous, i.e. the object is shared for fun, public shaming, etc. In a viral event, this wrapping tends to reaffirm itself by allowing little room for different and critical views (Ronson, 2016, p. 307). Moreover, most cases of viral content do not only involve forwarding of copies of the original content, but also the remixing and parodying of the original content. I will discuss the diverse kinds of signifying objects that can be a part of a viral outbreak in the next section.

However, it is not just the human agent as a social actor that plays a role in the forwarding and potential remixing of content: the mediating technology also plays a role. The affordances of the Web affect the potential publics of online content, the old dynamics of consent and the relation between publisher and author, as well as the affordances of digital objects themselves. Online, everyone can upload and edit content everywhere and at any time, without consent of either the original publisher or the referent (see chapter 4). As such, the Web itself easily affords every audience member to become a potential republisher of online content. Especially in cases of applications like social media, as discussed in chapter 5, the online architecture with its 'share' and 'retweet' buttons invites and simplifies the republication of information. Functions like the 'share' button thus propagate a certain distributive norm, while significantly accelerating the distributional affordances of the platform. Virality is therefore the result of a hybrid intentionality, in which the impact and role of the mediating technology differ per application.

7.4 Viral presence

In this section, I take a closer look at the presence of a viral reference. I will first discuss the incorporation of the reference in various signifying objects. Following this, I will discuss how these objects are spread through the network. Lastly, I will discuss the presence of the viral reference over time.

7.4.1 The object and its descendants

While virality is to a great degree a social phenomenon, it starts with the online accessibility of a particular signifying object: the initial object — 'patient zero' — which injected the particular reference into the online realm and from there let it go on an informational rampage. When the initial object is uploaded on the Web, it becomes open to the online affordances of easy transportation, multiplication, as well as easy editing (see section 4.2). As the object is picked up by users and mediating technologies, it is used as a base for the encoding of descendant objects like edited objects (remixes), hyperlinks, copies, search results, and feed objects (see figure 7.1). The edits can consist of anything from renaming to fundamentally altering the content. The descendant objects commonly consist of a newly stored piece of code and exist independently from the original object (although in the case of hyperlinks, with the removal of the original object, they break, but even in a broken form they still exist and can reveal some information).

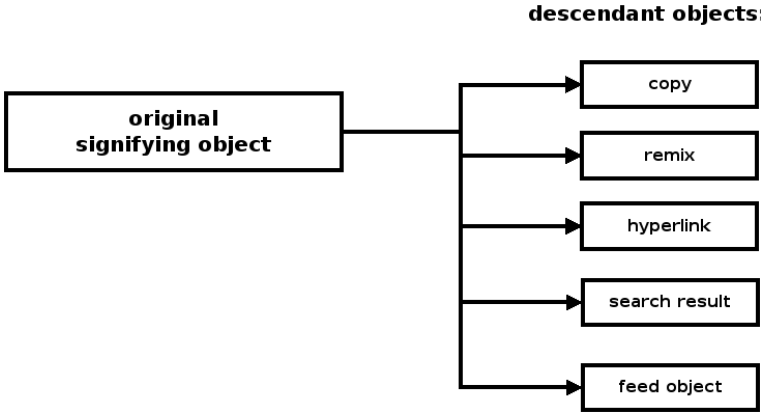


Figure 7.1: Signifying objects

These miscellaneous descendant objects have something in common: they share to a greater or lesser degree a certain reference with the previous object. For example, when the object 'Kneecam No.1' was picked up, framed as 'Technoviking', and turned into a variety of versions, the reference remained the typical appearance of the man with his distinctive moves. These descendant objects are thus new signifying objects sharing a similar reference, in this case the distinctive looking man dancing on techno music. It would therefore be more accurate not to talk of the virality of a particular signifying object, but of a particular reference.[13] The

[13]The popularity of remixing in relation to viral content is somewhat problematic with regard to the definition and 'identity' of viral content. For example, Shifman argues that *viral* content are objects that are distributed over the Web *without significant change* made to the original object (Shifman, 2012, 190). A viral signifying object "comprises *a single cultural unit* (such as a video, photo, or joke) that propagates in many copies" (Shifman, 2013, p. 55). However, this raises the question of when something can still be regarded as the same cultural unit, and

presence of the reference as expressed in the broad array of descendant objects is in a way the materialised footprint of a viral event.

The production of the descendant objects is part of a hybrid intentionality that touches deeply into the character of digital media. As Kelly points out: "every action you take on the Net or invoke on your computer requires a copy of something to be made. This peculiar superconductivity of copies spills out of the guts of computers into the culture of computers" (Kelly, 2007, p. 89). This conductivity of copies combined with the flexibility of digital objects gives rise to an online praxis and culture in which users become 'produsers' (users and producers) and "generate content by aggregating, mashing-up, (re)interpreting and distributing information" (Raffl *et al.*, 2011, p. 604). The praxis to create remixes, mash-ups, and the like, absorbs a reference into the online user culture. These practices are often accommodated by online applications that help users create signifying objects with just a few clicks. An example of this is `https://www.memegenerator.net` that enables a user to create an image macro with little effort. With this, the act of remixing is industrialised: it requires little know-how, effort or even creativity of the user producing the content.[14] By providing users with such tools, online applications not only facilitate the remixing and creation, but also bolster and co-shape the culture to do so. However, in the end, the user does remain the driving force in the creation of the majority of the descendant objects. As such, the quantitative and qualitative presence of the viral outbreak is dependent on a hybrid intentionality in which users play a key role.

---

when should it be considered a new unit. Moreover, given the importance of remixing in current Web culture, I argue that remixes should be seen as part of the viral event, and not as a subset, especially given the fact that they share a very particular reference. For this reason, combined with the fact that the remixes can have an equal, or even stronger, impact on the referent, I take the shared reference of objects to be the common denominator in a viral spread. In this context I find it important to distinguish this description from the concept of internet 'memes', because at first glance they may seem to have a similar character. Unfortunately, a 'meme' is a difficult concept because it is used in various manners, to the extent that some of the current use in internet culture is regarded as a 'hijacking' of the original idea, at least according to the originator of the idea, Dawkins (See Dawkins on the internet's hijacking of the word 'meme': `http://www.webcitation.org/6HzDGE9Go`, last accessed 28-03-2017). Due to the topic of this study, I will leave the academic discussion with regard to Dawkins' concept of 'memes' aside and will only focus on the meaning of the term according to its use in internet culture. On the Web, a 'meme' is best described as a group of online signifying objects that share a certain set of characteristics, style, and tone, and that were created, transformed and circulated by many users in awareness of the creation of similar objects by others (Shifman, 2013, p. 7-8). Examples of commonly used memes are for instance image macros. An image macro is an image on which a certain text is superimposed for a humorous effect. I argue thus that the difference between the shared character of memes and the shared character of viral content is that the shared character of memes lies in the form of the representation, while the shared character of viral content lies in the content of the reference. Hence, not all memes are viral content, and not all viral content is a meme.

[14]Such instant use technologies like the meme generator are not applauded by all online subcultures — and are sometimes even flat-out rejected because it turns an in-joke into a mainstream hit (cf. Miltner, 2014).

7.4.2 Viral spread

In a viral outbreak, descendant objects of the viral reference are spread over the Web. The spread differs per outbreak. Factors like the character of the content, the sender, the connectedness and popularity of the sender, the timing, and the context all play a role in the shape of the outbreak (Nahon & Hemsley, 2013; Jiang *et al.*, 2014). Also, it matters where the outbreak is triggered, for instance, on a social media platform or on a news site. The manner in which a viral reference is spread is therefore a combination of the (intertwined) social and technological nature of the Web, its applications and its users. In this subsection, I will discuss some of the main factors that shape the spread of a viral outbreak.

To begin with, the Web's architecture and networked character play a role. The high accessibility and conductivity of the Web affects the likelihood of a reference going viral by increasing the availability as well as the transmission speed of content. The consistent access combined with the hyperlinked and networked nature of the Web can potentially spread a reference worldwide in seconds. However, as described in section 4.5, the online audience is rarely, if ever, fully global. The spread of online information is generally centred in cultural subnetworks based on interest and/or background. As such, the social connections of users play a pivotal role in the viral information flow (Broxton *et al.*, 2013, p. 242). The distribution of viral content often takes place between peers who are networked within a particular application like a social media site, a forum or the like (Jiang *et al.*, 2014; Burgess, 2008). Within a certain interest network, information can spread relatively quickly due to the strong and/or overlapping ties between the users (Nahon & Hemsley, 2013, p. 31). When users are like-minded, they are more likely to value the same kind of content and maintain similar norms with regard to information sharing.

While viral information generally moves fast within a specific cultural subnetwork, it moves relatively slowly from one network to the other (Nahon & Hemsley, 2013, p. 31). This is especially the case when the networks are locked within applications that are online silos. While the content tends to spread rapidly within the silo, it has difficulty reaching users outside of this realm (see section 5.6). The technological environment can thus promote a viral spread in one direction, while hampering the spread in another direction.

In order for content to spread beyond a certain cultural subnetwork or a specific application, the networks need to be bridged. Some applications provide single-click tools for cross-application bridging. For instance, YouTube promotes cross-platform distribution by offering options to forward the content to email addresses or applications like Facebook. The increasing popularity of informational cross-references between platforms can increase the Web's informative conductivity and facilitate the spread of information (van Dijck, 2013, p. 101). However, while being single-click actions, users do need to make these bridges. Despite all the technological affordances and acceleration, a viral outbreak is in its core still a social event that heavily depends on user actions. Users thus play a pivotal role in initiating the forwarding and choosing when and how to insert content into

cultural networks.

There are three main (groups of) human agents who tend to play a role in the forwarding of content to an audience: (1) peers; (2) mass media; and (3) influentials (Cha *et al.*, 2012, p.993). Of the peers, it is commonly the weak ties (see section 5.6) who bridge different cultural networks and inject the viral content into a new cultural subnetwork (Nahon & Hemsley, 2013, p. 93). Traditional mass media can also play a role in the forwarding of viral content into new subnetworks. However, while mass media generally have large audiences, they have relatively little interaction: the user interaction on the online traditional media are often a 'mediated quasi-interaction' (Thompson, 2005, p. 33). Therefore the mechanisms of many traditional mass media are often too sluggish to participate in a viral spurt (Nahon & Hemsley, 2013, p. 54). Though, once a mass media picks a viral item up, the spread can increase quickly and easily reach new networks. However, the most influential agents for bridging distinct subnetworks are public or semipublic figures or entities like celebrities, politicians and local businesses. These are the influentials. Many of these people were initially not famous or public figures, but due to the affordances of the Web, and in particular with the mediation of social media, they managed to gain a strong public voice. Influentials play a leading role in the spread of information both by having a big audience (being popular) as well as enabling the connection between otherwise unconnected users and cultural networks (Cha *et al.*, 2012, p. 997). As such, they can reach a significant part of the general public with a relatively small group of broadcasters (Cha *et al.*, 2012, p. 994).

Next to the human agents, technological gatekeepers can play a significant role as actors in a viral outbreak. Algorithm-driven feeds (see section 5.5.2) and ranking mechanisms (see section 6.4) can push certain content forward. Especially given their often popularity-based evaluation mechanisms (like those discussed in chapters 5 and 6), they function as a catalyst in a viral outbreak; the more attention certain content receives, the more it will be brought under the attention of other users. As such, they can evoke a snowball effect that can result in a viral avalanche.

As the viral reference is spread by these different agents, we can see — depending on the spread — the occurrence of certain potential effects with regard to the presence of the viral reference. The most obvious effect, which occurs in all viral outbreaks (otherwise else it is doubtful that the content has actually gone viral), is the increase of the quantitative presence of the reference. With the mass sharing of a particular reference in various copies and possible remixes, the number of objects containing the particular reference are increased. The higher the presence of a particular reference, the higher the chance that a user is exposed to it at a certain point.

However, a viral outbreak also affects the forwarded reference on a qualitative level, which in turn affects its meaning. In case of a viral outbreak, I see two main effects that the event has on the meaning of the reference. First of all, the massive forwarding by human agents imbues the content with a certain social weight. The attention that the content receives as well as the forwarding, signals to people that

this is an object of interest (Nahon & Hemsley, 2013, p. 130). As such, being the topic of a viral outbreak, also imbues the reference with a qualitative presence. This qualitative presence is likely to go hand in hand with a change in the status of the reference: as the reference in a viral outbreak is handled by so many people, it may seem to have become not only public information, but a public good that everyone can share and remix.

Secondly, despite sharing a certain reference, all the descendant objects — even if they are exact copies — affect the meaning given to this reference: "every repetition of a sign involves an entirely 'new' semiotic process, allowing new semiotic modes and resources to be involved in the repetition process" (Varis & Blommaert, 2015, p. 36). To refer back to section 2.2.3, this means that despite the fact that two signifying objects may seem the same, like two digital copies, their signifying potential is somewhat different because they are two distinct objects, each embedded in their own context, with likely different interpreting users to which they signify something. With each copy, edit, and annotation, the object is recontextualised in a new situation, thereby affecting the meaning of the content. While these objects are tied by one common denominator, a certain reference, their particular context and mode of being can thus imbue this reference with a different meaning. As such, the reference is 'resemiotised' by being forwarded in descendant objects (cf. Iedema, 2001). This can complicate the interpretation of the meaning of the content, as well as whether, or which part of it, is real (Brown Jr, 2008). By combining content and/or moving it from one context into the other, new relations and interpretive settings come into being. This can even be the case, if the reference is something simple as a parody on Sesame Street's Bert. The 'Bert is evil' parody on Sesame's Street's Bert, in which the character of Bert is placed in compromising situations, suddenly appeared in an unexpected context (see figure 7.2):

Due to a hasty Google search by a company printing posters, this image of Bert alongside bin Laden was included on a protest poster used in Bangladesh. And so, Bangladeshi citizens protesting U.S. bombings in Afghanistan were waving signs that had Bert and bin Laden side-by-side—seemingly in cahoots. A Reuters photograph of the protest poster circulated via news outlets such as CNN and the New York Times, and the poster of Bert and bin Laden was seen by millions of confused Westerners. The image prompted a kind of hermeneutic fit from observers on message boards and various websites (Brown Jr, 2008).

Once references are placed together, their meaning can thus be reshaped in each others' context irrespective of the intentions of the publishers and users (Brown Jr, 2008). With the forwarding of the reference by myriad actors in a stream of descendant objects, the original publisher and context of the object can

[15]Original source: New York Times. Accessed through the Waybackmachine website, `https://web.archive.org/web/20020223032749/http://www.nytimes.com/learning/teachers/snapshot/student/20011015.html`, last accessed 14-03-2018. I blurred the faces of the protesters in order to protect their privacy.

Figure 7.2: Bert on protest poster[15]

even quickly disappear out of sight. As such, the affordances of the Web and its applications easily lead to collisions and combinations of references that can form new collaborations in the creation of meaning. In the case of a viral outbreak, users will therefore often interpret content that has already collided and mingled with various other objects and contexts, while this is not necessarily clear to the user. The reality of a specific viral reference is therefore fragile (Brown Jr, 2008).

7.4.3 Viral information life cycle

A viral outbreak is an event, something that happens. The viral reference therefore does not have a consistent presence in time. Instead, the outbreak has a certain life cycle with different phases in which the viral reference has a stronger or weaker presence. This life cycle consists of three or four potentially repeating steps: 1) the outbreak; 2) the decay; 3) the afterlife; and possibly a 4) revival phase (Nahon & Hemsley, 2013, p. 124). I will discuss these steps subsequently.

To start with the beginning of the viral outbreak. The outbreak is the point in time where the number of users that is confronted to a certain reference is sharply accelerating (Nahon & Hemsley, 2013, p. 25). The outbreak does not necessarily coincide with the time where the content is added to the Web. For instance, in the case of Technoviking the original content was uploaded in 2000 and it took until 2007 before the content went viral. The time of the outbreak can be connected to a certain naming, framing and editing actions of users (as was the case with for example Technoviking and Star Wars Kid), or on certain circumstances that raise a particular interest in the content, like in the Streisand case (see section 7.3). In the outbreak phase, the reference is widely spread with many copies and/or remixes. Also, due to the popularity of the content, it is likely to be placed prominently on websites and receive a high status in rankings and feeds. As such, the reference has a strong quantitative and qualitative presence in this phase.

After the outbreak follows the decay phase. The decay phase is the period in which there is a decrease in the speed and scope of the viral spread (Nahon & Hemsley, 2013, p. 125). This phase starts when there is a certain network saturation; many users in a particular network already took note of the content, and the content lost its novelty. In this phase, the content drops in feeds, rankings, and on websites. As such, especially the qualitative presence of the viral reference drops. As the content is tucked away and potentially even deleted, the quantitative presence can also drop. However, due to the storage-by-default of many of the online websites and applications, many of the copies and remixes of the viral reference are likely to maintain a lingering presence online (Nahon & Hemsley, 2013, p. 129) — just not in the centre of attention. A viral reference therefore likely maintains a relatively high presence compared to non-viral references.

As the viral reference remains lingering on the Web, it can easily receive a new round of attention. While in its afterlife a viral reference will for the majority depend on pull mechanisms for exposure, the availability of tools like search engines can easily place the reference at the centre of attention and contribute to a renewed interest in decaying viral content (Nahon & Hemsley, 2013, p. 131). Moreover, a (decayed) viral reference can be the topic of interest of researchers, journalists, etc.— as is the case here. When researchers and journalists publish about their findings, the reference may regain public interest. The interest of researchers and the like in viral content can lead to the archiving of the viral signifying objects or descriptions thereof in the public memory. An example of such an archive on viral content is the 'know your meme' website.[16]

Moreover, viral content in its afterlife can also be pushed to audiences by users who 'dump' their personal collection of interesting or comical signifying objects on sharing websites. Such dumping is a practice that we can see for instance on imgur.com. As one of the users commented on a dump: "These dumps are basically just recycled internet garbage. Sweet sweet internet garbage".[17] Such a content dump could recycle a viral outbreak.

With the many possible trigger mechanisms for the revival of interest in a viral reference, a new outbreak may always be around the corner (Nahon & Hemsley, 2013, p. 129). With that, the lifecycle of a viral event could be repeated ad infinitum.

7.5 Complications of the presented persona

The Web forms a fertile ground for viral outbreaks: signifying objects are easily and rapidly transported and multiplied, allowing them to reach massive audiences in mere seconds. Moreover, many online applications like social media offer simple publishing options and promote the sharing of content. The advance of viral content can be extremely fast thanks to these affordances of digital information. Online, a viral outbreak is therefore often just one click away: a signifying object

[16] http://knowyourmeme.com/, last accessed 20-09-2017.

[17] `https://imgur.com/gallery/7j5Dt`, last accessed on 10-02-2018.

can easily be forwarded from a relatively confined group to a massive audience (Ronson, 2016, p. 78).

When an individual finds herself to be the referent of a viral outbreak, this can pose some serious problems for her with regard to her informational persona. However, these problems have a multi-faceted character and may not always be clear at first sight. In the case of the Technoviking for instance, the content is generally perceived by the audience as positive and even with awe.[18] Despite this positive response, the subject experienced the virality as problematic. I argue that the explanation for this lies in the manner in which the viral reference constitutes an individual's informational persona *in relation to* the person of the referent. In this section, I therefore argue that problems raised for individuals by a viral outbreak run more deeply than merely the constitution of a negative portrayal of them (although in the case of negative content, the problems that the individual experiences are likely to be far more severe).

The most straightforward manner in which a viral reference affects an individual's informational persona is that by being excessively present, the viral content may easily outweigh other information. A viral outbreak gives rise to a multitude of relatively similar signifying objects and thereby casts a reference echo on the informational persona. Due to the excessive presence of the viral reference in the online information flows, users have a relatively high chance to encounter the viral reference — often even more than once. The presence of the viral reference can be so overwhelming that it drowns out other parts of the referent's informational persona and becomes the defining symbolisation of the referent on the Web (see e.g., Ronson, 2016, p. 264). As such, the online presence of a viral reference often results in a disproportional symbolisation of an individual; while the content may reflect a moment or minor aspect of a specific individual's life, the viral presence turns it into the main representation of the individual. This disproportional symbolisation of an individual by a viral reference is further enforced by the mechanisms of search engines that tend to prioritise the popular in the ranking of their search results (see chapter 6). If attributed to a certain name, the viral content is likely the top-ranked content for *any* individual that shares her name with the viral referent (Ronson, 2016, p. 264). The virality of a reference can therefore also affect others than the true referent.

The disproportionate nature of the viral reference is intensified by its often momentary and simple character. Generally a viral reference refers to a single moment in time, like taking a photograph or filming a particular event (see section 7.4.1). The resulting object is a 'singular sign'; it is the result of a recording at a unique moment and generally cannot be repeated in the same manner (Jappy, 2013, p. 87). As such, a viral reference thus often only reflects a *singular snapshot* in the life of an individual — if it is a realistic reflection at all. Moreover, given that viral content is often of a simple nature, the individual will be symbolised by what

[18]Fritsch' documentary about the Technoviking phenomenon gives a nice overview of the virality and the responses to the reference. Fritsch filmed the first video, 'Kneecam Nr. 1', and was the accused in the Technoviking court case. For the documentary and more information on Fritsch, see `http://www.technoviking.tv/subrealic.net/`, last accessed, 30-03-2019.

likely is a superficial glance of the individual, like a pose, a dance, a single sentence, etc. The two elements combined mean that a few seconds of an individual's life captured in a simple reflection can determine the manner in which the individual is symbolised online for years, if the reference goes viral. This impact of the viral reference on the informational persona will be more severe when there are few other signifying objects relating to the individual. Taking into account that most viral content features (previously) non-famous individuals (see section 7.4.1), the viral reference is highly likely to shape a significant part of the informational persona of those individuals because they are likely to have less other personal references online than public figures.

Over time, when the viral reference is in its afterlife, the impact of the viral content on the informational persona will decrease. However, in total, some of the viral objects are likely to remain intact at multiple locations and thereby continue to be a potentially significant part of the persona. The exact manner in which a viral reference's presence evolves in its afterlife, and whether it gets picked up again, depends on the hybrid intentionality of users and the websites that mediate the signifying objects. This differs per website. For example, content on 4chan.org disappears quite quickly due to the mechanisms that allow content to 'drop off' the site, while on other websites, the content can linger for decades (for instance, think about the seven years it took before the Technoviking video went viral). Content on social media is especially volatile in its presence over time: often old content is difficult to access, but due to several mechanisms, it can just as easily suddenly be picked up again and spread with high intensity (see section 5.5.2). Viral residue can increase again in its presence because the content can be picked up again, become a topic of court cases or trigger the interest of researchers. When this happens, it is likely that more objects containing the viral references are (again) added to the subject's informational persona, thereby strengthening the position of the viral reference as a symbolisation of the subject. With regard to such an interest induced revival, it is specifically the search engines that are likely to play a significant role because of the ease with which they allow users to retrieve information. Additionally, search engines can increase the identifiability of the referent by means of autocomplete (see section 6.4.1) and by combining different types and sources of information in one search result overview (see section 6.6).

Moreover, due to their high quantity and spread, viral signifying objects are difficult to control. Every user who has access to the object can make a copy and/or distribute the object further, thereby challenging the control of the original publisher. Often, the original users as well as the users republishing the content are unprofessional publishers and lack a code of conduct (Gregory & Losh, 2012) — or they do not consider any implications of the republishing for the referent in question. Also, the control over online signifying objects is further challenged by the potential response of users: attempts to control content may backfire and ignite a (new) viral wave (see section 7.3).

The combination of the broad spread and the lack of control severely hampers any attempt to correct or contextualise a viral reference. As such, it is almost impossible to undo the damage of a viral reference if for instance the content was

erroneous (Nahon & Hemsley, 2013; Hoskins, 2014). Even if an attempt would be made to spread a corrected version, it is unlikely that it would reach the same audience as its erroneous viral predecessor, especially if the correction does not go viral as well (and even if it would, it is questionable that the signifying objects reach the exact same audiences).

Meanwhile, the viral spread can claim a toll on the meaning of the reference. By reproducing a signifying object in another context, the content is resemiotisised (see section 7.4.2). With this resemotisation, the meaning of the reference is affected and can even be fully changed so that a skewed image of the referent arises. The potential minor and major edits to the signifying objects as result of the Web's remix culture can cause an even further resemiotisation of the reference. The Web's remix culture challenges the interpretation of the authenticity of objects and may lead to misinterpretations (Gregory & Losh, 2012). The result is that personal content that went viral can easily leave an impression on users in which they associate the subject with ideas, things and/or people with whom/which the individual herself has little or nothing to do.

Furthermore, the often pictorial character of a viral reference comes with its own set of consequences for the subject. Given that most viral references reflect the personal appearance of an individual, they can severely hamper an individual's ability to move anonymously in public space. The upside could be that in some cases changing one's appearance can be sufficient to distance oneself from a viral reference. However, when certain agents are adamant on identifying a certain individual that is portrayed in a particular signifying object, they can try to achieve this with techniques like facial recognition patterns if they have access to them. With the (future) developments in this field[19], it may only be a matter of time before users can search and retrieve information about individuals based on their face.

So far, all these problems are, beside their scale, not very different from the problems discussed in the previous chapters. However, I argue that virality has some critical implications for the informational persona that go beyond these problems. These implications result from the social side of the viral phenomenon.

With the massive forwarding, framing and viewing of the viral reference by a large number of users, the content becomes a sort of public good: the representation of the referent is appropriated by a public as part of a social phenomenon. As such, the representation of the individual is objectified and used as a means to an end — often for the purposes of entertainment (e.g., Star Wars Kid, Technoviking). The reference as part of the informational persona is 'hijacked' by the viral process in which the general public recontextualises the representation of the referent by naming, framing, remixing and changing the context. Most users feel like they can use, remix, and spread the image without consent of the subject. This objectification of the viral subject therefore goes hand in hand with a shift

[19]For example, see Shaun Walker, "Face recognition app taking Russia by storm may bring end to public anonymity", *The Guardian*, 2016. `https://www.theguardian.com/technology/2016/may/17/findface-face-recognition-app-end-public-anonymity-vkontakte`, last accessed 27-08-2019.

in norms; with the viral actions of many, the actions of remixing and sharing become the norm for that particular reference. Being objectified in such a manner can give rise to feelings of distress, loss of control, shame, and even depression in individuals (cf. Ronson, 2016). This adds to the distress already caused by the content, because often viral content portrays subjects in an undignified context where they are the object of ridicule, entertainment or public shaming.

Moreover, the process of viral forwarding generally places an emphasis on a particular aspect of the individual (which is often remarkable in one way or the other, see section 7.4.1). This emphasis is steadily ingrained in the signifying objects and steers users towards a certain way of interpreting, understanding and dealing with it by means of comments and edits. Particularly the emotional motivation for forwarding content can lead to a certain framing of a viral reference by setting a certain standard with regard to how to perceive the content. This standard is in turn strengthened with every forward within this framing. This public framing can even lead to the public shaming of people who voice doubts or critique about the framing (Ronson, 2016, p. 307). Dissenting users may therefore not voice their doubts or critique because they fear public critique, loss of popularity, and even public outrage.

As the heavy forwarding resemiotises the reference in a particular framing, the reference undergoes what I think can best be described as a certain 'symbolic wear'. The symbolic wear erodes the relation between the reference portrayed by the signifying object and the contextualised referent to whom it refers. While eroding the relation to the referent, the symbolic wear enhances the meaning attributed to the signifying object by the users who pass it on. Gradually the enhanced meaning shifts from a reference to a certain reality (e.g., man with a beard dancing on the street) to a reference to a concept ('Technoviking'); the narrative of the content becomes so enlarged and exaggerated that the viral subject is not only symbolised in a certain manner, but becomes a symbol for a particular character, way of acting, or a stereotype. As such, a viral reference places a strong stamp on an individual's informational persona by presenting the individual as a certain character portrait grounded in a particular unique representation of themselves. In this character portrait certain — often spectacular — aspects of the portrayed subject are amplified, thereby turning the personal reference into a *caricaturisation* of the individual. Especially the type of content that has a high viral potential is receptive to such caricaturisation; the content is often easy to digest, simple, pictorial and of such a nature that it invokes clear high-arousal emotions like anger, surprise and joy. This caricatured image is unlikely to correspond with the individual's view on her own identity.

However, not every viral outbreak will affect a subject equally. What plays an important role in the extent of the impact, is whether the audience takes the content to be representative for the individual. The relation between a viral reference and an individual's informational persona can differ: does the viral reference refer to a real or fictive referent? Take for example the Overly Attached Girlfriend. The Overly Attached Girlfriend video is likely understood by users to be an 'act': an individual that plays the role a fictional character, in this case an

obsessively attached girlfriend. They do not take the reference to be representative for the real person, but instead see a fictional character as the referent.[20] The *object* of ridicule and entertainment is thus not the subject as real life person, but the character she plays. However, this is closely tied to the credibility and style of the enactment. Contrary to the success of the Overly Attached Girlfriend-act, the Star Wars kid is an example of a less successful enactment. While the Star Wars Kid was performing an enactment of a jedi, his 'real' persona of a somewhat clumsy boy shimmered through. The result is that his performance was disrupted. Due to this disrupt, the Star Wars Kid as real individual behind the jedi became the referent of the viral outbreak. The viral content was thus publicly understood as representative for the real life individual's identity and character.

This is more complex in the case of viral references like Technoviking. While starting out as video footage of a street rave including a striking looking individual acting as himself and dancing on techno music, one may ague that with the symbolic wear of the content, the referent for the general public became more and more of a fictional character as 'Technoviking' became a concept and achieved a rock star-like status. With this, the caricaturisation may have effects in two opposite directions. On the one hand, the extreme caricaturisation may exacerbate the consequences for the individual, since the individual is reduced to a caricature while he still feels related to and represented by the content. On the other hand, the extremity of the caricaturisation may at the same time lessen the consequences, since the social response is not directed to a real life individual anymore, but to a fictional referent. What effect the caricaturisation in the end will have on referents and their informational persona — whether the viral content becomes part of the referent's real persona, or becomes a fictional persona on its own 'played' by the referent — is likely to vary per viral outbreak, per kind of content, and per referent.

The effects of online virality work through in a greater or lesser degree in the offline lives of the viral subjects. On the one hand, the degree to which the outbreak affects the offline life of the referent depends on how the referent experiences the outbreak and social responses to the content: does the referent feel that her own persona is objectified and caricatured (and maybe even violated), or does she feel that the outbreak is not about her but about a fictional referent? It is important to note here that even if the reference is clearly fictional, an individual can still feel that she is the referent of the content because her image or name is used. For example, a portrait picture of someone poorly pasted on extreme pornographic pictures, may clearly be fictional, but still the individual is likely to feel that she is the referent of the content and may feel that her persona is violated.[21] On the other hand, it matters to what extent and what part of the audiences understand

[20]This may explain why in this case the subject did not mind the virality of the reference. However, what also likely plays a role in the case of Overly Attached Girlfriend — as opposed to Technoviking, Star Wars Kid and Dog Poop Girl — is that the viral object is encoded online by the subject herself. Technoviking, Star Wars Kid and Dog Poop Girl were all posted by third parties without consent of the subject.

[21]See for example the Dutch 'Freek'-case, about a boy who's identity was used to create a fictional caricature, which deprived him of being able to construct his own persona in a meaningful manner. `https://www.kennisnet.nl/mijnkindonline/freek.html`, last accessed 19-06-2019.

a particular real life individual to be the referent of the viral outbreak. If the audiences do not take the reference to be representative for the real life individual, but instead understand a fictional character to be the referent, they are unlikely to respond towards the real life individual based on the content. The extent to which the audience attributes the viral reference to the real life subject will particularly matter with regard to the subject's main social circles (like friends, family, colleagues, neighbours and classmates) as these interact regularly with the individual. The more the individual's main social circles go along with the caricaturisation portrayed in the viral outbreak, the more severe the effect of the outbreak on the individual's life likely is.

The effects of a viral outbreak can have disastrous consequences for the subjects of the outbreak. The public objectification or outrage can cut deeply into the individual's life. In some cases, the virality can go as far as becoming an intentional witch hunt with the purpose of ruining someone's life (Dennis, 2008, p. 351). Especially in cases of negative responses like public shaming, the viral event can cause severe psychological distress and even lead to suicide.[22] The stress caused by a viral event is often intensified by the fact that moving past being the subject of a viral event can be very difficult (see the various cases in Ronson, 2016). Next to the severe initial stress, there is the always looming risk of a viral revival. With the significant presence of the viral reference in the information flow, even in its afterlife, the potential revival of the 'virus' may just be one click or one search string away. Especially the references that went viral as a result of public entertainment are relatively timeless due to their easy digestible content. However, not all viral references will be equally prone to revival. It is for instance questionable that Dog Poop Girl is suitable material for revival: this would require a second ignition of mass outrage for an already punished subject.

However, even without revival, the slumbering signifying objects can be problematic beyond the reference itself. Often, it is not just the reference that is stored, but also the public response to it. In the case of a humiliating or negatively perceived reference, these responses can be vicious and even constitute threats. What likely adds to this, is the disinhibition effect (see section 4.3.3), which lowers the barrier for people to leave hurtful comments compared to a face-to-face situation. For example, in case of the Star Wars Kid, the referent was confronted with vicious comments, and even comments telling him to commit suicide.[23] Moreover, due to the spatial and temporal affordances of online information, it becomes difficult for referents to escape these negative responses: connected to the Web, the content is consistently within access range of the referent and can also easily be pushed again towards her by feed mechanisms or other users. As such, referents may view the negative reactions 'over and over' (Campbell, 2005).

[22] See for example, Julian Robinson, " Italian woman commits suicide after sending taunting video of her having sex with new man to ex-boyfriend before footage goes viral on the internet", *Mail Online*, 2016. `http://www.dailymail.co.uk/news/article-3790966/Humiliated-Italian-woman-commits-suicide-sending-sex-tape-ex-boyfriend-taunt-uploads-Internet.html`, last accessed 21-03-2018.

[23] "10 years later, 'Star Wars Kid' speaks out", *Maclean's*, 2013. `https://www.macleans.ca/news/canada/10-years-later-the-star-wars-kid-speaks-out/`, last accessed 26-03-2019.

A lasting presence of online verbal aggression and public shaming can therefore have longterm consequences for the referent (Reid *et al.*, 2004, p. 243-244).

In sum, a viral reference places a disproportionally present, uncontrollable, essentially unrectifiable, caricatured stamp on the individual's informational persona. In this process, the subject is objectified and the target of a mass emotional framing. Virality causes a severe loss of control of the individual over a personal reference and with that over her informational persona and identity, which has become a public good. The viral reference can affect the public's view of the referent, while the resemiotisation and public annotations can deeply affect the referent's self-perception. However, as explained above, the extent of the impact of a viral reference depends on whether the audience understands the reference as referring to the real referent or to a fictional referent, as well as how the referent herself experiences the viral outbreak and its relation to her as individual. In a worst case scenario, the impact of a viral outbreak causes severe feelings of distress, shame and despair in the referent, and may even lead to suicide.

The technological mediation of the Web and its applications certainly express a particular intentionality in the process by accommodating, boosting and even inviting viral outbreaks due to the combination of the multiplication, transmission and editing affordances of online digital content combined with the push of the spectacular in many of the flow mechanisms of online applications (see chapters 5 and 6). However, the heart of virality problems lies in the social element of the viral outbreak: the public's use of and attitude with regard to the reference and the individual. Without the social mass motivation and use culture, the scope and severity of the impact on the symbolisation of the individual would be less far-reaching. Human intentionality therefore plays a crucial role in the coming into existence of the problems.

This role balance of the human and technological intentionality is somewhat reversed once the viral outbreak reaches its afterlife; in this stage the role of the mediating technology becomes the main factor in the construction of the presence of the reference due to its storage mechanisms. However, the chances of a viral revival are still closely tied to human actions: a viral revival will depend to a great degree on a human agent digging up a slumbering reference and pushing it back into popular culture. Yet, it is important to note that the Web provides a fruitful stage for this with its often longterm default storage combined with search engines which offer — especially popular — content ready-at-hand on the user's request. With a potential recurring outbreak just around the corner, the lingering online presence of a once viral reference is a perpetually looming sword of Damocles for viral subjects.

Chapter 8

Art. 17 GDPR

Contents

8.1 Introduction

1890. Two legal scholars became increasingly worried about the impact of the development of instantaneous photography and the increase in gossip press publications on 'the person' of the individual. In response, these scholars, Warren and Brandeis, wrote one of the most groundbreaking texts in Western legal history on the protection of the person. In their famous essay, *The Right to Privacy*, they state:

Recent inventions and business methods call attention to the next step which must be taken for the protection of the person (...). Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanismal devices threaten to make good the prediction that "what is whispered in the closet shall be proclaimed from the house-tops" (Warren & Brandeis, 1890, p. 195).

Their call for a better protection of 'the person' could just as easily have been uttered by a legal scholar in 2012, when the changing information landscape raised concerns with regard to the reach and retention of personal information published online. EU Justice Commissioner Reding stepped up to this challenge and argued that the changing technological landscape and corresponding business models required an update of the European data protection regime that was implemented in 1995. In this context, Reding argued that people should have better control over their personal information. She presented the idea to give this control shape in a 'right to be forgotten' and stated:

The Internet has an almost unlimited search and memory capacity. So even tiny scraps of personal information can have a huge impact, even years after they were shared or made public. The right to be forgotten will build on already existing rules to better cope with privacy risks online. It is the individual who should be in the best position to protect the privacy of their data by choosing whether or not to provide it. It is therefore important to empower EU citizens, particularly teenagers, to be in control of their own identity online.[1]

2019. We now have this 'right to be forgotten' in the form of art. 17 GDPR, named "Right to erasure ('right to be forgotten')". While this right is supposed to resolve the issues caused by assimilation of personal information by the Web, it is not clear yet whether it can, or to what extent. There seems to be a lack of a clear view on, and sometimes even a misconception of, the problems that art. 17 GDPR needs to address or what the right can do, or both. An example of this is the Drunken Pirate case, which I already touched upon in chapter 1. In my paper on this case, I have shown that art. 17 GDPR would not have been able to address

[1] Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*, `http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

the problem in this particular case, which ironically is often used as an example of why we need a 'right to be forgotten' (Korenhof, 2014). A better understanding of the problems was therefore vital in order to proceed to an evaluation of art. 17 GDPR as a means to address these. In the previous chapters, I therefore analysed how the Web affects the appearance of personal information to Web users and the manner in which problems were likely to emerge. The conclusion of these analyses is that the origins of the problems are less straightforward than an unlimited memory of the Web. In this, the Web turned out to indeed be a 'web': it is a tightly knit interplay between myriad online sources, agents, and applications that together shape the presence of online information — sometimes with problematic results. The informational persona can easily present an image of the referent that does not reflect her accurately or proportionally; the online persona can portray marginal elements as salient, it can reflect information in such a decontextualised manner that it is easily misinterpreted, it can lead into the persona becoming a public good, and/or it can undermine distance to the past. With the problems mechanisms clarified, the question is now: which (aspects) of these problems can be addressed by art. 17 GDPR?

This question brings me to the second challenge, which is the topic of this chapter. Assessing the right's functionality is not a simple matter of applying art. 17 GDPR to the cases, because the exact mechanisms of the right, and how we should understand these, are still a topic of discussion. In order to assess whether art. 17 GDPR can resolve the problems identified in this study, I will first need to construct a view on what art. 17 GDPR is, or could be, and how it works. Because much of the right still needs to take shape in practice, I will suggest an understanding that is at least partially instigated by the problem framework set out in this study: if art. 17 GDPR is supposed to resolve these problems, the right would benefit from being understood in relation to these. The detailed tracing of the roots of the problems provided by the problem analyses gives us a grip on how the problems come about, and can show us which elements can or should ideally be adjusted to resolve the issue. I therefore propose to construct an understanding of art. 17 GDPR and its merits to address the problems that is partially built upon what we know of them. The main part of constructing an understanding of art. 17 GDPR is, due to its legal character, necessarily rooted in its legal text and context. Although the right is tightly connected to and dependent on other provisions of the GDPR, I will approach art. 17 GDPR in this chapter with a focus on the text of the article itself and how the article's specific functionality can interfere with information processing on a practical level. The reason for this is that the goal of this study is to assess specifically what art. 17 GDPR can bring to the table when it comes to resolving certain (pivotal) elements in the emergence of the problems. The process of assessing art. 17 GDPR's problem-solving potential will thus require an exploration of the possibilities offered by the practical workings and restrictions of art. 17 GDPR in relation to the problems, that, in turn, is used to assess the functionality of the right in the specific cases discussed chapters 4 to 7. To structure this assessment, I propose to perform it in two steps: first, I will identify the possibilities offered by the legal workings and restrictions of art. 17

GDPR to address problems, as well as potential pitfalls, so that we have a clear baseline to work with. Secondly, with this baseline in hand, I will combine all of the previous chapters into a bigger picture and sketch a particular understanding of art. 17 GDPR as a means to address the identified problems, which, in turn, I will apply to the four explored cases (Web, social media, search engines, and virality). In order to keep a clear overview, I will split these two steps over two chapters by discussing step one in this chapter, and step two in chapter 9.

In this chapter, I will take thus take a closer look at art. 17 GDPR. In order to investigate the workings of art. 17 GDPR, I will focus on its text, because the text of the article is certain and we have to make do with how it is formulated. I will analyse the mechanisms of art. 17 GDPR by means of close reading of the article's text, complemented by case law where needed. I will combine this with the knowledge that we have of the problems to give some direction to the elements that I explore. Also, I will examine the right's name. By having a closer look at the article itself, I investigate what the right itself can tell us about its goals and functionality. The goal of this chapter is to determine how the right works, and explore where its strengths and weaknesses lie. The analyses of the previous chapters will serve here as a theoretical framework. I will not discuss art. 17 GDPR's full legal status as being embedded in our broad juridical system, because this is of little help to answer the question of whether art. 17 GDPR itself is a suitable means to address the problems on a practical level. Rather, this chapter will provide an interpretation of art. 17 GDPR that aims to clarify how the right itself works in the context of online information processing. This interpretation will be used in the next chapter to elaborate on the relation between the right and the problems identified in chapters 4, 5, 6, and 7, and assess to what extent art. 17 GDPR is a viable means to address these problems.

Lastly, for clarity's sake, I print the full text of art. 17 GDPR on the next page.

Art. 17 Right to erasure ('right to be forgotten')

1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:

a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;

b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1)[2], or point (a) of Article 9(2)[3], and where there is no other legal ground for the processing;

c) the data subject objects to the processing pursuant to Article 21(1)[4]and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2)[5];

d) the personal data have been unlawfully processed;

e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;

f) the personal data have been collected in relation to the offer of information society services referred to in Article 8(1)[6].

2. Where the controller has made the personal data public and is obliged pursuant to paragraph 1 to erase the personal data, the controller, taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures, to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data.

3. Paragraphs 1 and 2 shall not apply to the extent that processing is necessary:

a) for exercising the right of freedom of expression and information;

b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

c) for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2)[7]as well as Article 9(3)[8];

d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1)[9]in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing; or

e) for the establishment, exercise or defence of legal claims.

[2]The data subject gave consent for the processing of her personal information.

[3]The data subject gave consent for the processing of special categories of her personal information, like sexual preference and health information

[4]The data subject has the right to object to the processing based on grounds relating to her particular situation.

[5]The data subject has the right to object to the processing of her personal information for marketing purposes.

[6]The services are offered to a child.

[7]The processing of personal health information is allowed for the health care purposes (h) or for the benefit of the public interest in the area of public health.

[9]This processing needs to be done with appropriate safeguards like pseudonymisation and

8.2 The mechanisms

In order to evaluate the viability of art. 17 GDPR to address problems, we need to know what art. 17 GDPR is and how it works. However, it is not an easy task to clarify this: what 'the right to erasure ('right to be forgotten')' actually is, is still a matter of discussion and legal development. Given the novelty of the GDPR, there is still little case law on the GDPR itself. Old case law can be of help here at some points. Art. 17 GDPR has many characteristics of the old right to request deletion of personal information under the DPD (see e.g., the analysis of art. 17 GDPR by van Hoboken (2013)), and there is quite some case law on the right to deletion on the Web. Nonetheless, I am wary to interpret the right too strongly under DPD case law. The reason for this, is that developments in ICT technologies, especially those online, have been ongoing at high speed, thereby changing the scope, scale and character of the manner in which the Web mediates personal information. The applications on and of the Web in 1995 (the year the DPD was adopted) are very different from those in 2019. One of the reasons for the development of the GDPR was to do justice to these new technological developments and 'update' the laws. Understanding the GDPR solely in the light of DPD case law may constrain some of its concepts and applications too narrowly to views that see to relatively outdated situations. This can undermine the GDPR's potential to effectively deal with contemporary technologies. This does not mean that I will not look at case law altogether; I will touch upon relevant case law, especially in relation to the balance of interests. However, in order to get a grip on the problem-solving potential of art. 17 GDPR, I will primarily approach the article by looking at its foundation: its functional mechanisms.

In this section, I will therefore discuss art. 17 GDPR per functional element. I trace the required practical steps for invoking and applying art. 17 GDPR, discuss what art. 17 GDPR does and does not do, and point out the cases where its interpretation is unclear or potentially problematic. The elements that I will address are art. 17 GDPR's rationale, the pivotal actors in the GDPR, namely the data subject and the controller, the material scope of art. 17 GDPR, its territorial scope, its target (signifying objects), the concept of erasure, the role of the subject, the grounds on which the right can be invoked, and lastly, the exceptions to the execution of the right.

8.2.1 The rationale

Because technological developments can give rise to new ways of collecting, disclosing and disseminating information, an increased legal protection of personal information can be deemed necessary. We see this view already expressed by Warren and Brandeis, and also for example by the European Court of Human Rights (ECtHR), which argues that "increased vigilance in protecting private life is necessary to contend with new communication technologies which make it possible

anonymisation, in order to respect the data minimisation principle.

to store and reproduce personal data"[10]. The GDPR is such a response to the rise of new technologies. The European Commission recognised that technological developments brought challenges for the protection of personal information. These challenges, combined with the great degree of variation in the manner in which the DPD was implemented throughout Europe, motivated the Commission to develop the GDPR. EU Commissioner Reding gives several reasons for its introduction:

17 years ago less than 1% of Europeans used the internet. Today, vast amounts of personal data are transferred and exchanged, across continents and around the globe in fractions of seconds. The protection of personal data is a fundamental right for all Europeans, but citizens do not always feel in full control of their personal data. My proposals will help build trust in online services because people will be better informed about their rights and in more control of their information. The reform will accomplish this while making life easier and less costly for businesses. A strong, clear and uniform legal framework at EU level will help to unleash the potential of the Digital Single Market and foster economic growth, innovation and job creation.[11]

The goal attributed by the European Commission specifically to art. 17 GDPR is to help individuals manage the risks of sharing their personal information online by allowing them to have the information erased "if there are no legitimate grounds for retaining it"[12]. This protection is considered especially important with regard to information that is processed based on consent, and where this consent is given by the individual when she was a child (recital 65 GDPR). Even when the individual has grown up, she still has the right to have content erased to which she consented as a child (recital 65 GDPR).

The rationale attributed to the GDPR and specifically to art. 17 GDPR by the legislator give some foothold on how to approach the right, but overall remain rather fuzzy. The reasons for introducing the GDPR given in the press release quoted above focus on different points (i.e., control for individuals, the establishment of trust in online interactions, making life easier, and less costly for businesses), which at times likely move in different directions that may even be incompatible. With regard to art. 17 GDPR, the legislator made clear the right is there to strengthen the position of individuals with regard to their personal information, especially if this information refers to them as a child.[13] While this gives some direction on how to understand the right, this still leaves much in the open. The rationale underlying art. 17 GDPR seems thus somewhat underdeveloped.

[10] ECtHR, 25-06-2004, application no. 59320/00 (*Von Hannover v. Germany*), §70.

[11] European Commission press release, Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses, `http://europa.eu/rapid/press-release_IP-12-46_en.htm`, last accessed 19-07-2018.

[12] Ibid.

[13] Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*, `http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

8.2.2 Data subject and controller (and processor)

Two pivotal actors in the GDPR are the 'data subject' and the 'controller'. Given their importance, I will first discuss their respective definitions and how they relate to the terminology used in the previous chapters, before I delve more deeply into the mechanisms of art. 17 GDPR. Lastly, it is important to acknowledge the role of a potential processor. However, as I will explain, I will not consider this actor further in the rest of the evaluation of art. 17 GDPR.

Data subject The data subject is "an identifiable natural person (...) who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (art. 4(1) GDPR). As such, the data subject is the person to which information refers, and which I have called 'referent' and 'subject' in my analyses in the previous chapters. Given the meaning of the word 'data' as discussed in chapter 2, I have chosen not to use the term 'data subject' in the previous chapters, but only in these last chapters where I refer to the referent in her role as a legal subject.

Controller In the previous chapters, I have on many occasions used the term 'controller', albeit often with an extension, like 'medium controller'. I have used this term in line with the GDPR. The GDPR defines the controller as "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data" (art. 4(7) GDPR). Sometimes the control over the processing of personal information lies in the hands of multiple controllers at the same time. In this case, we speak of 'joint controllers'. Agents are a 'joint controller' when they "jointly determine the purposes and means of processing" (art. 26(1) GDPR). The joint controllers need to make their arrangement clear to the data subject (art. 26(2) GDPR). In the case of a joint controller, the data subject can invoke her rights against each of the controllers (art. 26(3) GDPR).

Processor Lastly, I will briefly touch upon the concept of 'processor'. The processor of information is the "natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller" (art. 4(8) GDPR). The purposes and means of the information processing performed by the processor is determined by the controllers. WP 29 explains that the distinction between controllers and processors serves to allocate responsibility: "[t]he distinction between 'controller' and 'processor' mostly serves to distinguish between those involved that are responsible as controller(s) and those that are only acting on their behalf"[14].

Because the controllers determine the purposes and the means of the processing, they are the relevant actors in art. 17 GDPR requests and the corresponding

[14] WP 29, Opinion 1/2010 on the concepts of "controller" and "processor".

balance of interests. I will therefore treat the processors and controllers as one actor in the evaluation of art. 17 GDPR, and will not consider the role of processor separately.

8.2.3 Material scope

The first question with regard to the workings of art 17 GDPR, is its scope: to what does it apply? Art. 2 GDPR defines the material scope of the GDPR and states that it "applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system". The main conditions for the application of the GDPR, are set by the combination of two main elements, namely the (1) *processing* of (2) *personal data*.

'Processing' is broadly defined in the GDPR and entails "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction" (art. 4(2) GDPR). In short, doing *anything* with personal information on a computer falls within the scope of the GDPR (cf. Voigt & Von dem Bussche, 2017). Given that the GDPR's concept of 'personal data' is used as a parameter for this study (see section 1.2.1), we can conclude that the cases discussed in the previous chapters clearly fall within the material scope of the GDPR. I will therefore not discuss the fringes of these definitions.

However, there are some exceptions to the material scope of the GDPR that are relevant for the applicability of art. 17 GDPR to online personal information. I will discuss these here.

8.2.3.1 2(2)(a): processing falls outside of Union law

Art. 2(2)(a): **"This Regulation does not apply to the processing of personal data (...) in the course of an activity which falls outside the scope of Union law"**

With art. 2(2)(a) GDPR, the processing of personal information for activities that fall outside the scope of European Union law, is placed outside the scope of the GDPR. Recital 16 gives the processing of personal information for national security as an example of this restriction to the scope of the GDPR.

8.2.3.2 2(2)(b) and (d): border security, public safety and prosecution of criminal offences

The GDPR does not apply to the processing of personal information by governments in order to battle criminality, for the purpose of public safety, or in order to develop and execute border, asylum and immigration security policies. This

is codified in exemptions (b) and (d). As these exemptions show some family resemblance (they are both aimed at maintaining order and safety), I will discuss them together.

Art. 2(2)(b): **"This Regulation does not apply to the processing of personal data (...) by the Member States when carrying out activities which fall within the scope of Chapter 2 of Title V of the TEU;"**

Art. 2(2)(d): **"This Regulation does not apply to the processing of personal data (...) by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security"**

These exemptions see to a very particular setting: the personal information is processed by authorities for specific purposes conform exemption (b) or (d). Exemption (b) refers to chapter 2 of Title V of the Treaty on the Functioning of the European Union (TEU) concerns freedom, security and justice with regard to national border policies. Exemption (d) connects to Directive 2016/680, which sees to the protection the processing of personal information by authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, as well as to the free movement of this information.

Because both exemptions see on very particular occurrences of information processing, which for the majority are not likely to result in publicly accessible information about particular individuals, and are regulated by separate instruments, I will leave the discussion of the publication of personal information by national and international authorities for the aforementioned purposes outside the scope of this study.

8.2.3.3 2(2)(c): household exemption

Art 2(2)(c) GDPR places the processing of personal information purely for a personal or household activity outside the scope of the GDPR. This is the 'household exemption'. This exemption enables individuals to shape and retain their personal tertiary memory without limitations on the processing, as long as the processing has "no connection to a professional or commercial activity" (recital 18 GDPR). Recital 18 explains: "Personal or household activities could include correspondence and the holding of addresses, or social networking and online activity undertaken within the context of such activities". However, there is another restriction to the processing in order for the content to fall under the household exemption: the information should not be shared with larger groups beyond the family and personal friends of the person who processes the information. A noteworthy case in this context is the Lindqvist ruling by the

CJEU.[15] In this case, the court decided that publishing personal information about volunteers of a church community on a website does not fall under household use. While Lindqvist did not have any commercial intentions, her information processing was considered to fall outside the scope of household use because the content on a publicly accessible web page can be viewed by an indefinite number of people.[16] In order for an online signifying object to qualify for household use, the access to the object should thus be limited. However, the exact scope of what is considered 'limited' enough to fall under the household exemption is still unclear. This is especially challenging with regard to social media. WP 29 advised that in the case of social media, the application of the household exemption should require a limitation of the audience to a *self-selected* set of contacts.[17] Additionally, in order to qualify for the household exemption, there is a limit to the number of self-selected contacts. WP 29 states: "A high number of contacts could be an indication that the household exception does not apply and therefore that the user would be considered a data controller"[18]. Unfortunately, the WP 29 does not give an indication of what would qualify as a high number. I will discuss the details of the application of the household exemption to social media further in section 9.4.1.2).

It is important to remark that the scope of the household exemption is a topic of discussion. The CJEU's decided in the preliminary ruling in the *František Ryneš* case that processing of information by a surveillance camera attached by a private person to his own house for the safety of his own property and family, but directed partially to the public space around his house, did not fall under 'purely household use'.[19] The court opted for a narrow understanding of the household exemption that only covers the processing of personal information *purely* as a part of household use. It argued that: "To the extent that video surveillance such as that at issue in the main proceedings covers, even partially, a public space and is accordingly directed outwards from the private setting of the person processing the data in that manner, it cannot be regarded as an activity which is a purely 'personal or household' activity"[20]. This narrow interpretation would, as I read it, also place any post on social referring to a data subject outside of the household exemption as this processing transcends the context of purely personal activities due to the interactive platform structure on which the processing takes place.[21] The narrow interpretation of the household exemption in this ruling led to critique and gave rise to various views on, and even applications of, the household exemption (Ausloos,

[15]CJEU, 06-11-2003, C-101/01, ECLI:EU:C:2003:596 (*Criminal proceedings against Bodil Lindqvist*).

[16]Ibid., §47.

[17]WP 29, Opinion 5/2009 on online social networking, p. 5-6.

[18]Ibid., p. 6.

[19]CJEU, 11-12-2014, C-212/13, ECLI:EU:C:2014:2428 (*František Ryneš*).

[20]Ibid., §33

[21]To the extent that I read the ruling wrong and it is actually about the content that covers the public space, this would mean that every holiday picture in a private photo album would not be covered by the household exemption if the photo was taken in the public space and shows strangers tagging along in the background. I find this unlikely, as this is exactly the kind of content that I think the household exemption should protect.

2018, p. 149-152).

For the purposes of this study, I leave aside this particular narrow focus of the household exemption because it is contested and would rule out so much of the personal information processing from protection under the household exemption that the exemption seems to become rather void. Instead, I will follow the wider scope as expressed by WP 29 in their opinion on social media. The reason for this is, first of all, that this opinion seems in line with recital 18 of the GDPR which explicitly mentions that social networking can fall under the household exemption. If it is even possible to use social media in such a manner that it would fall under the narrow interpretation of the household exemption, such use would likely defy the purpose of social media use. The wider scope of the household exemption ties in better to the contemporary use of online media. Secondly, I focus on the wider scope of the household exemption because the wider its scope is, the more it will hinder a successful application of art. 17 GDPR. If over time the household exemption will evolve in a more narrow direction, this will be to the advantage of data subjects who wish to invoke art. 17 GDPR.

The household exemption, at least in the form as set out by WP 29 in their opinion on social media, tells us that art. 17 GDPR is not meant to lead to a full-fledged erasure of personal information in any given context. It is thus not a right that allows data subjects to indiscriminately exercise control over what information a specific other has about them. Instead, it targets information processing in the public and semipublic realm (if a large enough number of users has access to the content), as well as organisation controlled (though potentially not publicly shared) information collections. The core of the right thus lies in giving users some control over their societal informational persona.

8.2.4 Territorial scope

The GDPR only applies to the processing of personal information that falls within its territorial scope. This scope is listed in art. 3 GDPR. There are two main triggers for the GDPR's territorial scope: (1) the establishment of a controller or processor on EU territory (art. 3(1) GDPR), and (2) the targeting of EU data subjects (art. 3(2) GDPR).[22] Additionally, the GDPR applies to the processing of controllers who by virtue of public international law fall under the application of the laws of one of the Member States (art. 3(3) GDPR). However, as the first two triggers require the most explanation and are also the most interesting in the context of this study, I will restrict my discussion of these triggers to the first two.

The first GDPR-trigger, the presence of a relevant establishment of a controller or processor on EU territory, is codified in art. 3(1) GDPR, which states: "This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not". The last part of art. 3(1) GDPR, "regardless of whether the processing takes place in the Union or

[22]European Data Protection Board, Guidelines 3/2018 on the territorial scope of the GDPR (Article 3) - Version for public consultation, p. 3.

not" gives this provision an extraterritorial dimension. The core of the provision depends on whether the controller or processor has (1) an establishment in the EU, and (2) the personal information is processed in the context of the activities of this establishment (van Alsenoy, 2017). In recital 22, 'establishment' is explained as "the effective and real exercise of activity through stable arrangements". The European Data Protection Board concludes that the threshold to speak of an establishment through stable arrangements is quite low: "the presence of one single employee or agent of the non-EU entity may be sufficient to constitute a stable arrangement if that employee or agent acts with a sufficient degree of stability"[23].

Next to having an establishment in the EU, the processing of the personal information needs to be performed in the context of this establishment in order for the controller or processor to fall under the GDPR. Thus far, this context is interpreted in a broad manner and covers a direct as well as an indirect link between the processing activities and the establishment (van Alsenoy, 2017, p. 84). Although trialled under the DPD, it is relevant to note that in the *Google Spain* case, the CJEU applied a broad interpretation of 'in the context of activities' by arguing that the processing of search results is linked to the selling of advertisement space.[24] The CJEU therefore concluded that the "processing of personal data is carried out in the context of the activities of an establishment of the controller on the territory of a Member State, (...) when the operator of a search engine sets up in a Member State a branch or subsidiary which is intended to promote and sell advertising space offered by that engine and which orientates its activity towards the inhabitants of that Member State"[25]. The broad interpretations of these concepts seem to have been adopted in art. 3(1) GDPR , thereby imbuing the GDPR with a significant extraterritorial scope (van Alsenoy, 2017, p. 84).

While the territorial scope on the level of the potential controllers that fall under the GDPR is wide, the CJEU in a recent case applied art. 17 GDPR's erasure only to information processing within a limited territorial scope. In the *Google v. CNIL* case, the CJEU ruled that "where a search engine operator grants a request for de-referencing (...) that operator is not required to carry out that de-referencing on all versions of its search engine, but on the versions of that search engine corresponding to all the Member States".[26] Moreover, with regard to the delisting of content within the EU, the CJEU indicates that the scope of delisting needs to be balanced against the freedom of information, and that the balancing in a specific case may differ per Member State.[27] However, the CJEU does not fully rule out the possibility of a global application of erasure. The court states: "Lastly, it should be emphasised that, (...) EU law does not currently require that the de-referencing granted concern all versions of the search engine in question, it also does not prohibit such a practice. Accordingly, a supervisory or judicial authority of a

[23]European Data Protection Board, Guidelines 3/2018 on the territorial scope of the GDPR (Article 3) - Version for public consultation, p. 5.

[24]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §55-57.

[25]Ibid., §60.

[26]CJEU, 24-09-2019, C-507/17, ECLI:EU:C:2019:772 (*Google v. CNIL*), §73.

[27]Ibid., §67.

Member State remains competent to weigh up, in the light of national standards of protection of fundamental rights (...) to order, where appropriate, the operator of that search engine to carry out a de-referencing concerning all versions of that search engine"[28]. The CJEU thus leaves it up to the Member States to decide whether the erasure should also entail the delisting of search results in non-EU domains. How this will take shape in practice remains to be seen.

Secondly, art. 3(2) GDPR provides the GDPR with an even stronger extraterritorial scope. This provision places the processing of personal information of EU data subjects by controllers who do not have an establishment in the EU within the scope of the GDPR when this processing is performed for "(a) the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or (b) the monitoring of their behaviour as far as their behaviour takes place within the Union" (art. 3(2) GDPR). The core of art. 3(2)(b) GDPR lies in the *targeting* of EU data subjects (De Hert & Czerniawski, 2016, p. 238). Both (a) and (b) revolve around the targeting of EU subjects *in* the EU (thus excluding, for example, the local activities of EU data subjects when they are on holiday in the US) (De Hert & Czerniawski, 2016, p. 238). Starting with art. 3(2)(a) GDPR. This provision provides the GDPR with a significant territorial scope as many controllers outside the European Union explicitly offer their goods and services to subjects in the European Union. This is the case for many controllers in the United States, but also elsewhere. Take for example the Asia-based AliExpress.com. By offering on their site information translated into multiple European languages like Dutch and German, combined with shipping options to European countries and the display of prices in Euros, it seems clear that the controller is offering its goods specifically (also) to subjects in the European Union. By doing so, websites like AliExpress.com fall under the scope of art. 3(2)(a) GDPR.[29][30]

Art. 3(2)(b) sees on the behavioural monitoring of data subjects. If a controller intentionally targets EU data subjects by monitoring them, the processing of the personal information of these subjects of the controller is likely to fall within the territorial scope of the GDPR. De Hert and Czerniawski capture the logic of art. 3(2) GDPR in the following rationale: "you might be targeted by EU law only if you target" (De Hert & Czerniawski, 2016, p. 238).

Despite the GDPR's broad territorial scope[31], a part of the controllers will

[28]CJEU, 24-09-2019, C-507/17, ECLI:EU:C:2019:772 (*Google v. CNIL*), §72.

[29]Merely presenting online goods and services in the language of a particular EU country is insufficient to conclude that a controller offers goods or services to EU subjects. However, the combination of "factors such as the use of a language or a currency generally used in one or more Member States with the possibility of ordering goods and services in that other language, or the mentioning of customers or users who are in the Union, may make it apparent that the controller envisages offering goods or services to data subjects in the Union"(recital 23).

[30]One can raise some serious questions with regard to the manner in which art. 3(2)(a) establishes the GDPR's extraterritorial scope. However, discussing this lies outside the scope of this study. I would like to refer readers interested in this topic to the article *Expanding the European data protection scope beyond territory: Article 3 of the General Data Protection Regulation in its wider context* by De Hert & Czerniawski (2016).

[31]The manner in which the GDPR establishes an extraterritorial scope has been widely

fall outside of it. In these cases, data subjects will be unable to invoke their right to erasure. However, the GDPR may still have some influence in these cases. I can imagine that certain corporate or governmental controllers who fall outside the scope of the GDPR, may be sensitive to EU regulation for political or economic reasons. This will likely be different for amateur controllers located outside the EU who post content online. With the global and open publishing character of the Web, such amateur controllers may easily decontextualise content by (re)publishing it in a different (cultural) context. To the extent that these controllers fall outside the GDPR's territorial scope, their signifying objects will remain untouched by art. 17 GDPR. As such, the limits to the GDPR's territorial scope may put a halt to the ability of art. 17 GDPR to address some of the identified problems.

8.2.5 The targeting of signifying objects

By allowing individuals to demand from a controller the erasure of personal information relating to them, art. 17 GDPR is focused on a concrete target: already existing personal signifying objects in the hands of a particular controller. The targeted content can consist of anything ranging from a single signifying object to a huge dataset.

Koops as well as Ausloos point out that by focusing on existing content, art. 17 GDPR works *ex post* and can only be used after the information is processed — used — for something (Koops, 2011; Ausloos, 2012, p. 243). By focusing only on existing content, art. 17 GDPR does not prevent the creation of new content, but it can prevent the creation of new descendant objects.

8.2.5.1 Targeting solely the controllers of descendant objects

If we look at the problems identified in this study, many issues are caused by third parties who made use of the affordances of online objects and created descendants of an original object (e.g. remixes, hyperlinks, search results, copies). One of the crucial questions with regard to the efficacy of art. 17 GDPR therefore is, whether it can be invoked *solely* against the controller of a descendant object (this could even be the original controller who processes the information for new goals), while leaving the original processing of the object untampered with.

This question played an important role in the *Google Spain* case: can a data subject require the erasure of a search result while the original object to which the result links, remains accessible on its source website?[32] Although technically the case was not trialled under the GDPR, but under the DPD, I take this case to be a good indication of the working scope of art. 17 GDPR as the GDPR replaces the DPD. With this case as precedent, art. 17 GDPR offers the removal

criticised (see e.g., Svantesson, 2015; Kuner, 2015). However, discussing this critique falls outside the scope of this study.

[32]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

of descendant objects created by a search engine, while leaving the original content intact. Descendant object controllers can therefore be treated as independent targets for art. 17 GDPR's application.

Targeting controllers of solely descendant object(s) is a valuable asset of the right for solving some of the identified problems; in many cases the creation of the descendant object is the cause of the problem by establishing a decontextualised, overrevealed, or disproportional presence of a particular aspect of an individual's informational persona. We can see this clearly in certain cases of search results, where a search engine highlights and recontextualises the presence of a marginal and/or outdated piece of information. Take for example the cases discussed in chapter 6, where several individuals who were interviewed about testicular cancer wanted to have the search results referring to these interviews removed. While the interviews on the BBC page remain intact, and also retrievable by means of search engines for those looking for information on testicular cancer, the highlighted presence of the testicular cancer reference in the informational persona of the specific individual is undone. By solely erasing such descendant objects, certain cases of salience, decontextualisation, and overrevealing may be addressed by art. 17 GDPR. I will discuss this further in the next chapter.

8.2.6 Erasure

Art. 17 GDPR gives individuals — under certain circumstances — "the right to obtain from the controller the erasure of personal data concerning him or her without undue delay" (art. 17 GDPR). The main functionality of the right is thus *erasure*. The application of the right takes place ex nunc and without undue delay.

While at face value the 'erasure' of information may seem clear enough, the concept of 'erasure' is on closer inspection one of the most ambiguous aspects of art. 17 GDPR.[33] What, exactly, qualifies as *erasure* in art. 17 GDPR? Does erasure necessarily equal the complete deletion of the object, or a 'mere' removal from view?

First of all, complete erasure in the digital milieu can be technically difficult to realise when you do not intend to physically destroy any parts of the device. Deleting a signifying object by means of for instance clicking 'delete' in a menu, often does not mean that the binary code of the digital object is gone, but rather that the index through which it can be accessed in the device's storage is destroyed. With the right tools (e.g., TestDisk[34]), these seemingly deleted objects can often be retrieved. The thoroughness of erasure required by art. 17 GDPR is a topic of speculation.

[33] Typically enough, the term 'erasure' seems to have been chosen in order to integrate "the right to have the processing restricted in certain cases, avoiding the ambiguous terminology 'blocking'". See Explanatory Memorandum, 52012PC0011, November 2012, `https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52012PC0011&from=EN`, last accessed 13-07-2019.

[34] `https://git.cgsecurity.org/cgit/testdisk/`, last accessed 07-11-2019.

The *Google Spain* case can provide valuable guidance here.[35] When we take this case as a precedent for art. 17 GDPR, we need to conclude that erasure does not have to entail a full erasure from the database. In this case, the CJEU ordered the removal of a particular search result *in relation to a specific name search*. This erasure of certain search results *in response to* a particular query does not entail the complete erasure of the record from the search engine's database, but only the removal of a particular descendant object of the record in the list of results following a specific query. The record is thus merely blocked from being displayed as search result in particular search queries. It remains available in the database and can still be retrieved with other queries. In these cases, art. 17 GDPR is applied not in the form of erasure, but as an active blocking from appearing. When intervening with search results, art. 17 GDPR is therefore also referred to as 'right to be delisted' (see e.g., Peguera, 2015; de Mars & O'Callaghan, 2016). Moreover, as discussed in section 8.2.4, in the recent *Google v. CNIL* case the CJEU applied a delisting scope that was limited to search results shown in EU accessible versions of the search engine.[36] I will discuss the implications of a limited delisting scope in section 9.5.1.3.

While the ruling in the *Google Spain* case was based on the at the time still active DPD, I take it as a good indication for what should be considered as 'erasure' under art. 17 GDPR.[37] Moreover, as art. 12(b) DPD is the predecessor of art, 17 GDPR[38], combined with the above mentioned ruling, suggests that we should understand 'erasure' of art. 17 GDPR to be shorthand for some of the actions referred to in art. 12(b) DPD that prevent access to particular information: erasure and blocking. The opposite, understanding 'erasure' not as a shorthand for more actions, but only as a definitive erasure action, not only makes little sense given the history and context of the right, but additionally it would turn art. 17 GDPR in a relatively static right that can only 'erase' in the narrow sense of the word. This would not be to the advantage of its problem-solving ability, which I will discuss in detail in the next chapter.

However, this relatively open interpretation of erasure does leave us with the question of what is considered to be sufficient erasure under art. 17 GDPR. I argue that if we want to make the most of art. 17 GDPR, the manner in which erasure

[35] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[36] CJEU, 24-09-2019, C-507/17, ECLI:EU:C:2019:772 (*Google v. CNIL*), §74.

[37] There is a debate whether, and to what extent the CJEU evolved art. 12(b) and 14 DPD into a new right in this ruling (Bartolini & Siry, 2016; Politou *et al.*, 2018b). As the DPD is already receded, I shall not go into this question as it lies outside the scope of this study.

[38] See Explanatory Memorandum, 52012PC0011, November 2012, which states: "Article 17 provides the data subject's right to be forgotten and to erasure. It further elaborates and specifies the right of erasure provided for in Article 12(b) of Directive 95/46/EC and provides the conditions of the right to be forgotten, including the obligation of the controller which has made the personal data public to inform third parties on the data subject's request to erase any links to, or copy or replication of that personal data. It also integrates the right to have the processing restricted in certain cases, avoiding the ambiguous terminology 'blocking'.", `https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52012PC0011&from=EN`, last accessed 13-07-2019.

should be applied differs highly per case. This is where my problem analysis of chapters 4 to 7 comes in. In the next chapter I will discuss the degree of erasure needed to address the various identified problems.

Lastly, it is important to mention that the controller plays an important role in deciding what 'erasure' entails in a particular case. When an individual files an erasure request at the address of a particular controller, it is initially up to that controller to grant or deny a request for erasure, and how to realise the erasure. This places significant discretionary power in the hands of the controller (Koops, 2011, p. 240). The subject and the controller could even maintain different ideas on what 'erasure' should entail in a specific situation. An exemplary case of this was discovered by Schrems, when he found out that Facebook merely made invisible much of the information that users thought they erased, while it retained the information in its database.[39] If users do not understand the manner in which controllers apply 'erasure' correctly, they may easily end up in a situation where their information is retained beyond their knowledge, and with that, beyond their possible control through the GDPR. As such, the concept of 'erasure' shows the difficult marriage of a legal text with the practical reality of the Web.

8.2.6.1 The tail of erasure

As already pointed out, if an individual successfully invokes her right to erasure, the controller of the content does not only need to erase the information, but also "taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures, to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data" (art. 17(2) GDPR).

I understand paragraph (2) of art. 17 GDPR as an attempt to account for the typical affordances of digital information, especially online, that allow third parties to easily reproduce and edit content and continue its processing elsewhere. The ideal consequence of paragraph (2) would be that with the erasure of the source, all its descendant objects would also disappear. However, given the affordances of online objects and the praxis of online information flows, it is questionable how realistic this is. Once online, virtually everyone can copy, store, hyperlink to, and distribute the object. While these actions do leave some traces in the log files of websites (indirectly in the case of hyperlinks), it is likely a burdensome task for controllers to figure out who all the third parties are and what they have done (i.e., they might have downloaded or copied the object, but that does not indicate whether they have further disseminated or published the content) — that is, if the controller retains her log files, else tracing all the descendant objects would become virtually impossible. Moreover, the controller only needs to inform the third parties of the erasure request. This does not mean that these parties will comply with the request (or even worse, it may trigger a counter reaction). This happened with the BBC cases. Here, the removal of certain search results following

[39] See `http://europe-v-facebook.org/EN/Data_Pool/data_pool.html`, last accessed 25-10-18.

a name search, prompted a BBC editor to publish a list of all the removed URLs (see section 6.3). Also, these third parties may process the information on different legal grounds. This may allow these third parties to continue processing, while the original object (or set of objects) is erased. The court case of the Techoviking is an example of this. Here the removal of the original objects is legally obliged, while the third party processing in the form of remixes is left untouched.[40] Taken everything together, I expect the cases of successful erasure of all the descendant objects of the targeted object(s) as a result of the application of art. 17(2) GDPR, to be scarce.

8.2.7 Subject driven

One of the core elements of art. 17 GDPR is the data subject. When tracing the required practical steps for the application of art. 17 GDPR, the first thing to consider is the fact that art. 17 GDPR is a subjective right; the data subject needs to invoke her right to erasure. This is a possible weakness of art. 17 GDPR.

First of all, the data subject needs to be aware of the fact that under certain circumstances she can have a right to erasure. Taking a bit of an optimistic stance that controllers do meet their obligation to inform data subjects of their right to erasure (art. 13(2)(b) and art. 14(2)(c) GDPR), I assume for the purposes of this study that many data subjects have been informed about their rights somewhere during their onli*f*e. If data subjects are not aware of their right to erasure, an awareness campaign will be needed to bring art. 17 GDPR to the attention of data subjects if it is to resolve any of the identified problems at all.

Secondly, the data subject needs to be aware of the problematic processing before she can invoke her right. Koops, as well as Ausloos, therefore point out that due to this dependency on the subject, the right is likely to be invoked in many cases only after the problems of the information processing already transpired (Koops, 2011; Ausloos, 2012).

Another possible set of complications is that in order to have content erased under art. 17 GDPR, the data subject needs to invoke her right to erasure against a particular controller (or in the case of a joint controller, the data subject can invoke art. 17 GDPR against each of the controllers (art. 26(3) GDPR)). In order to invoke art. 17 GDPR against a particular controller, the data subject will need to identify and contact this controller. Contacting the controller can be difficult; not all websites list how to contact the controller and Internet Service Providers are not likely to give up the contact details of their clients easily (I will discuss this further in section 9.3.1.1). Moreover, it can be difficult to even determine who the controller is of a particular signifying object, especially in cases where there are multiple entities involved. However, the joint controller construction provides a solution here. I will discuss this in relation to the cases in section 9.4.1.1.

Next to invoking art. 17 GDPR against a particular controller, the subject will also need to base her request on one of the grounds listed in art. 17(1)(a)-(f) GDPR. I will discuss these grounds per ground in subsection 8.2.8. For now, it is

[40] Landgericht Berlin, 30-05-2013, 27 O 632/12.

important to note that except in the case of ground (b), the withdrawal of consent, I expect that the data subject will need to argue why she wants that particular content erased. This places a part of the burden of substantiating an art. 17 GDPR claim with the data subject.

In the application of art. 17 GDPR, the capabilities of individuals to locate content, determine who the controller is, and argue convincingly why it should be removed, thus play a significant role. This becomes even more pressing when the subject is looking to have multiple signifying objects removed. The subject needs to track all the content and request its removal. This may turn out to be a burdensome quest for the average person. Especially the highly networked character of the Web combined with the digital character of online objects affords a quick and broad spread of content and allows for a variety of processing actions and controllers (see chapter 7). However, the subject can receive support to invoke her erasure requests. For instance, a dedicated NGO could file an injunction based on art. 79 (the right to an effective judicial remedy against a controller or processor) in conjunction with 80 GDPR (the representation of data subjects) and act on behalf of the data subject. Such support would ease the burden on the subject.

Additionally, with paragraph 2, art. 17 GDPR attempts to alleviate the burden caused by the reproductive affordances of online content. Art. 17(2) GDPR stipulates that the controller takes reasonable steps to contact third parties who further processed the content, and to inform them that the subject wants to have it erased. As such, this paragraph could prompt the erasure of all the descendant objects together with the object from which they derive. However, it seems unlikely that this will work so easily in practice due to the dynamic character of the online world (I will discuss this in more detail in section 8.2.6). If the controller takes reasonable steps but does not succeed in contacting these third parties, the burden shifts back to the data subject. The result is that, when an individual wants many or even all descendant objects erased (as for instance a subject of a viral case may want), she will need to (1) be able to track all objects that contain a particular reference, (2) be able to get a hold of contact information of the controllers of these objects, and (3) file a request for erasure with every controller. This is a Herculean task (see e.g., Korenhof & Koops, 2013). However, I expect that in many cases, it might be sufficient to have the object removed from the most obvious and most-visited sites to resolve the current problem. The prospective problem of a potential re-viralisation will remain due to the remaining objects, but with the reduced presence of the objects, this risk is also reduced (I will discuss this in the next chapter where I delve into the application of art. 17 GDPR on the cases discussed in chapters 4 to 7).

8.2.8 Grounds

The data subject can invoke art. 17 GDPR only on particular grounds. The grounds tell us something about the when and why of art. 17 GDPR. In this subsection I will therefore take a closer look at these grounds and investigate what they reveal about art. 17 GDPR.

8.2.8.1 Ground (a)

"the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed"

Ground (a) strongly ties in with one of the main principles of the GDPR, namely purpose limitation (art. 5 (1)(b) GDPR). While the principle of purpose limitation is a matter for research on its own, I will briefly outline its rationale and use as ground for the application of art. 17 GDPR.

The purpose limitation principle existed before the GDPR: we can find it in art. 5(b) of the Council of Europe Convention No. 108 on data protection, art. 8(2) European Union Charter of Fundamental Rights, and it took further shape in art. 6(1)(b) of the DPD. The purpose limitation principle restricts the processing of personal information to processing that is in line with specified purposes. The information shall be "not further processed in a manner that is incompatible with those purposes" (art. 5(1)(b) GDPR). This means that the controller cannot process the information for new purposes from the original processing without making sure that she has a corresponding new legitimating ground.

The purpose limitation principle has two building blocks: "personal data must be collected for 'specified, explicit and legitimate' purposes (purpose specification) and not be 'further processed in a way incompatible' with those purposes (compatible use)".[41] It is important here to briefly point out that some further compatible processing is necessary to even be able to comply with the GDPR: 'storage' as well as 'erasure' are also forms of processing (art. 4(2) GDPR). Erasure of information is, if it is no longer needed for the purposes for which it was collected, therefore a compatible form of further processing.

The purpose limitation should "prevent the use of individuals' personal data in a way (or for further purposes) that they might find unexpected, inappropriate or otherwise objectionable"[42]. However, the purpose limitation principle is not only of importance for the protection of individuals, but also serves the 'rule of law' by ensuring that "the powers of the state (and possibly of civil actors) are constrained for the protection of rights and liberties" (Brouwer, 2011, p. 276). As 'information is power', the purpose limitation principle is an important restriction that curtails the power imbalance that may rise between citizens and the institutions that can aggregate massive amounts of information about them (Brouwer, 2011, p. 280). I therefore argue that this function of the purpose limitation principle is especially important with regard to online corporations and institutions that collect online user information.

In the GDPR, the purpose limitation principle has become more flexible. Art. 6(4) GDPR builds a certain flexibility into the purpose limitation principle by allowing further processing for a different purpose if the new purpose is compatible with the purpose for which the information was initially collected.[43] Moerel and

[41] WP 29, Opinion 03/2013 on purpose limitation, p. 3.

[42] Ibid., p. 11.

[43] This has been a major cause of worry for the WP 29. WP 29 even recommended the deletion

Prins argue that with the increased use of the collection of personal information, not as a byproduct, but as a goal to offer services to users, this more flexible approach of the GDPR towards the purpose limitation principle is a necessity to cope with future technological developments (Moerel & Prins, 2015).[44] Yet, this relaxing of the purpose limitation principle may make it difficult for users to be able to successfully invoke art. 17 GDPR on ground (a). This is even more difficult when they target information on the Web, because this information is processed for a thousand-and-one purposes, and stored by default (Ambrose & Ausloos, 2013, p. 7). With regard to content that is published online, we can imagine that while the initial stated purpose of a particular publication was to inform the general public, the ongoing retention of the content on the Web may be done under the flag of archiving purposes. Jones and Ausloos therefore conclude that the online information processing practices are "rendering the purpose limitation principle quite toothless in practice" (Ambrose & Ausloos, 2013, p. 7).

While ground (a) ties in with the purpose limitation principle, it is not an exact replica of the principle. Its difference from the purpose limitation principle may give it (a bit) more practical usability in the online world. The core of ground (a) here, I argue, lies in the combination of the fact that art. 17 GDPR is subject driven and has a focus on information that is '*no longer necessary*'. I see ground (a) therefore as a tool that individuals can use to argue that the ongoing processing of particular personal information referring to them is no longer necessary. This would suggest that something has changed in the situation, preferences or wishes of the individual, in the situation of the controller, or in the context of the information, which overturns the initial purpose of the information collection. In this sense, ground (a) may be best seen not as a purpose *limitation*, but as a purpose *expiration*[45]. As such, art. 17 GDPR serves as a tool for individuals to challenge which processing purposes are, at a given point in time, still representative for their lives and self-presentation. In some online cases, ground (a) may therefore be a useful ground. An individual could successfully invoke art. 17 GDPR on ground (a) if she can make a case that the controller is processing the information in a manner that, after a certain lapse of time, does not serve the purposes of the processing anymore, likely as a result of changes in the life and interests of the individual, like a change in professional career, the ending or establishment of a relation, changes in one's outlook on life, etc.

of paragraph 4 (WP 29, Opinion 03/2013 on purpose limitation, p. 41.).

[44]Moerel and Prins argue that, instead of the purpose limitation principle, the base for legitimate processing should be the legitimate interest. They therefore advocate to replace the purpose limitation test with a legitimate interest test (Moerel & Prins, 2015).

[45]Ausloos also refers to 'purpose expiration'. However, we differ in our use of the concept: he relates it to three main principles in art. 5(1) GDPR: purpose limitation, data minimisation, and storage limitation (Ausloos, 2018, p. 163), while I relate it specifically only to art. 17(1)(a) GDPR in relation to a change of circumstances.

8.2.8.2 Ground (b)

"the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing"

Art. 6(1)(a) allows controllers to process personal information based on the consent of the individual, and art. 9(2)(a) allows such processing based on consent with regard to special, and therefore more strongly protected, categories of personal information. Personal information is considered to be 'special' if it reveals "racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation" (art. 9(1) GDPR). Consent is a relevant ground for information processing under the GDPR, especially with regard to online content. However, once a data subject gave consent, this does not mean that her consent is indefinite: art. 7(3) GDPR allows data subjects to withdraw their consent.

The manner in which consent takes shape in the online environment, has been target of critique. As Koops argues, online consent "is largely theoretical and has no practical meaning" (Koops, 2014). Here he refers to the online reality of 'consent': users easily give consent to access services without spending much thought to what they agree to. Users tend to give consent without even reading what they are consenting to (Obar & Oeldorf-Hirsch, 2018). While users easily give their consent online, controllers gladly make use of this and use consent as a ground for all sorts of processing that they would not be allowed to do otherwise. Consent is therefore often misused in order to bypass limitations on information processing (see e.g., Moerel & Prins, 2015; Politou *et al.*, 2018b).

While consent is easily given, it can thus also be withdrawn at any time due to art. 7(3) GDPR. If consent is withdrawn, the controller loses her legitimate ground for further processing. The withdrawal of consent itself does not automatically affect the processing of personal information that has been performed prior to the withdrawal (art. 7(3) GDPR). However, the controller does need a legitimate ground to be allowed to keep the information in storage or otherwise process the data and needs to inform the data subject of this new lawful basis conform the transparency principle required by art. 13 and 14 GDPR.[46] This is where art. 17(1)(b) GDPR comes in. By invoking art. 17 GDPR an individual clearly withdraws her consent and requests the immediate erasure of the information relating to her, as well as requiring the controller to inform third parties of her request for erasure.

While as a ground, ground (b) is hardly surprising as the controller loses her legitimate ground for processing by the withdrawal of consent, it does have a practical and empowering use because the erasure of personal information is not a necessary consequence of the withdrawal of consent on the basis of Art.

[46] WP 29, Guidelines on consent under Regulation 2016/679, p. 22-23.

7(3) GDPR. Ground (b) thus empowers users to maintain a certain discretionary power over their personal information, even when, at a certain point in time, they consented to the processing of their information. Jones and Ausloos therefore argue that as a right to erasure, art. 17 GDPR's function is to shift the power between the controller and the data subject on the level of consent (Ambrose & Ausloos, 2013, p. 15). In this capacity, art. 17 GDPR "might help to cure the shortcomings of the [online] consent regime" (Ambrose & Ausloos, 2013, p. 12). With ground (b), art. 17 GDPR thus is a means to 'undo' the processing based on online consent.

8.2.8.3 Ground (c)

"the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2)"

Art. 21(1) and (2) GDPR provides individuals with the right to object to the processing of their personal information under certain circumstances. I will discuss these here. I will not delve into the details of this right, as it requires a research as extensive as the one being performed here for art. 17 GDPR, but I will point out its main elements.

Art. 21(1) GDPR states: "The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims". This provision gives an individual the right to object to the processing of her personal information *on grounds relating to her particular situation.* The wording suggests that, when someone objects to the processing of her personal information, she needs to have a specific motivation for this objection in relation to her particular situation. We can think of examples like publications referring to happy marriages, while the data subject is now divorced, sport reports on the subject's performance in a soccer match, while she may now be hindered from playing any sports due to permanent injuries as result of a traffic accident, or a blog about a summer camp portraying a family's experiences, while shortly after the interview two of the children died in an accident. A hypothetical example that I worked out in detail with Koops, is the case of Agnes, who was born as 'Andrew', and wanted to be free of her past as Andrew (Korenhof & Koops, 2013). In all of these cases, I think it is not difficult to imagine that a data subject does not want to be confronted with the reference anymore, either herself directly, or likely even worse, indirectly by others who are unaware of her current situation ("I saw nice pictures of your wedding, how is your husband?", "I saw your family photos of the summer camp, you have lovely children!"). This is thus a relevant ground with regard to online

information and the focus of this study because it covers many situations where personal information is made publicly available, like blogs, the publication of news articles, websites that inform users about professionals, etc. However, the question in relation to invoking art. 17 GDPR on this ground is if, and to what extent, the subject will have to submit arguments why she invokes her right to object. If the data subject has to give arguments, and may even have to show evidence, this can prove to be a significant (especially emotional) obstacle for the subject to invoke art. 21(1) GDPR. Having to provide arguments in order to invoke their right to object, might therefore create some chilling effect on subjects due to which they are hesitant to share details of their current situation.

However, the core of the burden of proof with regard to art. 21(1) GDPR seems to lie with the controller.[47] Recital 69 states that "[i]t should be for the controller to demonstrate that its compelling legitimate interest overrides the interests or the fundamental rights and freedoms of the data subject". With regard to the used vocabulary, WP 29 pointed out that the use of the word 'compelling' in the article suggests a relatively high threshold for controllers to be allowed to continue their processing despite the individual's objection to the processing of her information.[48] Leaving aside where the burden of proof lies (this may very well be a research on its own), art. 21(1) GDPR shows that a certain significance is given to an individual's particular situation in weighing the balance of interests. As such, the life and person of the individual, and possible changes in these, become an important factor in the application of art. 17 GDPR.

Art. 21 GDPR (2) is a powerful right with regard to cases where information is processed for direct marketing purposes; on this level, the right to object is unconditional as the article does not leave room for the controller to make a case for the continuation of the processing.[49] The question is to what extent art. 21(2) GDPR is effective as a ground with regard to the focus of this study. This depends on the relation that the GDPR and following case law require between the user of personal information and direct marketing. If the relation needs to be one-on-one, i.e. solely the processing of personal information of the subject to market directly to the subject herself, e.g., by means of profiling, then this ground has little use for the researched cases. However, if the scope also covers the use of personal information for direct marketing to others, it may be of use in cases where personal information of a particular individual is used to market directly to other users (see section 5.3). To give an example, when a photo of a social media user is used in an advertisement to her friends, she (if she finds out that this has happened) could try to invoke art. 17 GDPR on ground (c) in relation to art. 21(2) GDPR. If this indeed would fall within the scope of art. 21(2) GDPR, the unconditional character of art. 21(2) GDPR would turn art. 17 GDPR into a strong instrument to enforce erasure in cases where personal information is used for direct marketing to others.

[47] WP 29, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, p. 19.

[48] Ibid., p. 19.

[49] Ibid., p. 19.

In sum, ground (c) allows individuals to request the blocking of the processing of their personal information based on her personal situation, or in case the processing is performed for direct marketing. With this ground, art. 17 GDPR provides individuals with a strong tool to influence the content of their informational persona, based on (a) its relation to their lives, or (b) attempts to seduce them or others to do particular purchases. Ground (c) in conjunction with art. 21(1) GDPR acknowledges the impact of the construction of the informational persona on the life of the individual, and empowers the individual to address, under certain circumstances, this impact by means of erasure. In addition, the value hierarchy expressed in art. 21 GDPR suggests us how to approach art. 17 GDPR in this context. Comparing paragraph (1) and (2) of art. 21 GDPR, we see that individuals receive a stronger measure of control over their personal information when it comes to the use of this information for direct marketing. Advertisement is thus given a weak position in the balance of interests corresponding to ground (c). From this, we may also derive some clues with regard to the general relative weight that should be given to some other forms of commercially motivated processing when diverse interests need to be balanced in an art. 17 GDPR case; the EU legislator seems to give relatively little weight to the interest of controllers that process personal information purely for commercial purposes compared to the protection granted to the data subject. Thus while controllers have the right to conduct a business, I expect that their interest will in many cases lose out against the interests of the data subject invoking art. 17 GDPR if the controller processes the subject's personal information *purely* for commercial purposes.

8.2.8.4 Grounds (d) and (e)

(d): "the personal data have been unlawfully processed"

(e): "the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject"

Grounds (d) and (e) seem rather obvious: when the controller is unlawfully processing personal information (d) or is legally obliged to erase the information (e), the individual has a right to have the content erased. The added value of art. 17 GDPR in these cases, lies in the subject driven character of the right; these grounds support the aim to give control to the user by giving her an active role in realising GDPR compliance. Moreover, when individuals invoke art. 17 GDPR on these grounds, this could work as a signal to the national Data Protection Authority (DPA) that something is wrong or could even relieve a part of their burden if the controller responds to such individual requests by adjusting her processing altogether. Unfortunately, this is the optimistic view. The pessimistic view entails that since the controller is already disobeying the law, it seems unlikely that she would suddenly comply with the data subject's erasure request. In this case, the data subject will need to turn to the DPA for help.

Given the legal orientation of these grounds, they seem to be primarily of use to legally proficient individuals. To successfully invoke art. 17 GDPR on one of these grounds, the individuals thus need to be more legally proficient than merely being aware of their right to erasure.

8.2.8.5 Ground (f)

"the personal data have been collected in relation to the offer of information society services referred to in Article 8(1)"

This ground deals with information that is collected from children by information society services (art. 8(1) GDPR). Information society services are defined in art. 1(2) of Directive 98/48/EC[50] as "any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services". The collection of personal information by these services is lawful if the child is at least 16 years old (art. 8(1) GDPR). If the child is younger, the parent or guardian of the child needs to consent to this processing (art. 8(1) GDPR).

The rationale underlying ground (f) is that a child is not fully aware of the consequences and the risks when she gives up personal information (recital 38 GDPR). Moreover, the protection "should, in particular, apply to the use of personal data of children for the purposes of marketing or creating personality or user profiles and the collection of personal data with regard to children when using services offered directly to a child" (recital 38 GDPR). The withdrawal of consent in combination with an erasure request is therefore considered to be of particular importance when the consent was given by a child (recital 65 GDPR). Even as an adult, the individual can request the erasure of information that she shared when she was a child (recital 65 GDPR).

In her speech, Reding underlined the importance of the right to erasure especially for teenagers as a means to give them a certain degree of control over their identity construction.[51] This added importance is justified, because adolescents are likely to experiment with their identity online as they try to figure out who they are and what they want in life, and are in a phase of life in which they tend to undergo significant changes. If an individual makes use of social media throughout her teens and puberty (i.e., children are allowed to use Facebook from the age of 13 with consent of their guardians), she may easily leave an information trail that she later regrets. This may result in cases of 19-year olds who try to find a job and therefore want to have certain photos removed that they uploaded when they were 15, and of 16 year olds who are embarrassed by content that they uploaded (with parental consent) when they were 13 years old.

[50] They were initially defined in art 1(2) of Directive 98/34/EC, and later amended by Directive 98/48/EC.

[51] Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*`http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

Ground (f) mainly seems to be there to make sure that individuals can distance themselves from youthful mistakes and actions they may regret later in life. Especially as people tend to experiment in their youth and develop themselves strongly between youth and adulthood, the protection given by art. 17 GDPR to the information that people share as children can be an asset of the right.

8.2.9 Exceptions

There are several exceptions to art. 17 GDPR, that may prevent the subject from having her personal information erased. In this subsection, I will take a closer look at these exceptions and investigate what they can tell us about art. 17 GDPR and the problems that it is supposed to resolve.

Please note: in this section I only focus on the rationale of the exceptions and what they can tell us about art. 17 GDPR. I will not yet discuss the meaning of these exceptions for the particular cases discussed in chapters 4 to 7. This will be done in chapter 9.

8.2.9.1 Exception (a)

"for exercising the right of freedom of expression and information"

For the online world, given its character as a public communication network, the most important exception to art. 17 GDPR is exception (a); when the processing is necessary for exercising the right to freedom of expression and information. This exception is further strengthened by art. 85 GDPR which requires Member States to reconcile data protection rights with the right to freedom of expression and information by law. Art 85(2) GDPR states: "For processing carried out for journalistic purposes or the purpose of academic artistic or literary expression, Member States shall provide for exemptions or derogations from (...) Chapter III (rights of the data subject) (...) if they are necessary to reconcile the right to the protection of personal data with the freedom of expression and information".

While the value of art. 17 GDPR for individuals and society still needs to be decided (although I hope to offer some help for this with this study), the value of the freedom of expression has long been acknowledged and established. Freedom of expression, especially with regard to the press and political speech, receives significant protection in the European Union.[52]

Legal literature acknowledges several purposes of the right to freedom of expression and information. The right is considered of great importance for the autonomy of the individual, it is seen as a means that may help to discover what is true, it can help realise progress, it is an important instrument for democratic and political decision-making, it can help people to socially bond, and it can function as a safety-valve against feelings of repression (cf. Nieuwenhuis, 2011). However, legal scholars differ in their views on purposes that are served by the

[52]See e.g., ECtHR, 26-04-1979, application no. 6538/74 (*Sunday Times v. UK*), ECtHR, 08-07-1986, application no. 9815/82 (*Lingens v. Austria*).

freedom of expression and information, as well as the respective priority that these purposes should be given (Nieuwenhuis, 2011, p. 21-22). The views that people hold on these purposes and their respective weight, shape the weight and role that is given to freedom of expression and information in particular cases. Due to the differences in views on the purposes and priorities with regard to the freedom of expression and information, I cannot provide the reader with a clear and encompassing view that has a universal foothold. Nevertheless, given the fact that the GDPR is European legislation, the manner in which the values are prioritised by the ECtHR can provide a useful guidance when it comes to the application of art. 17 GDPR. The ECtHR prioritises the value that the freedom of expression and information has for the public interest in a democratic society over other values (Nieuwenhuis, 2011, p. 289). This value embodies a sliding scale with regard to the content of the expression; while political speech and press are heavily protected, commercial speech and publications that mainly serve for amusement are given lesser protection (Nieuwenhuis, 2011, p. 292).[53]

Unsurprisingly, art. 17 GDPR as the 'new kid on the block' that moves in the same playing field as the freedom of expression and information, is often described as a threat to the right to freedom of expression and information, especially by scholars based in the United States (see e.g., Rosen, 2011; Larson III, 2013; Fazlioglu, 2013). According to Rosen, art. 17 GDPR even poses "the biggest threat to free speech on the Internet in the coming decade" (Rosen, 2011, p. 88). Understanding art. 17 GDPR and the freedom of expression and information as opposites is to a certain extent logical, given that it is listed as an exception. However, I suggest a more nuanced view in which both can co-exist and actually support the reasons for each other's existence.

First and foremost, the right to freedom of expression is not absolute in the EU. In general, the right to freedom of expression and information has been curbed by inter alia limits with regard to discriminatory and pornographic material (Nieuwenhuis, 2011). And while especially journalists and politicians enjoy a strong protection to their freedom of expression, also their expression can be restricted when the interests of others outweigh the interests of the expresser and the general public with regard to the information. The freedom of expression of politicians has for example been restricted when their expression was considered discriminatory.[54] With regard to the press, we can find cases in which the freedom of expression has been restricted when it made a disproportional impact on the private lives of individuals (even when these are public figures)[55]. The freedom of expression, at least in a European context, is thus a right with a broad scale of nuances in weight and applications. I argue that it is from this perspective, that we should approach the balance between art. 17 GDPR and the freedom of

[53]See e.g., ECtHR, 05-05-1979, application no. 7805/77 (*X and Church of Scientology v. Sweden*), ECtHR 24-05-2005, application no. 59320/00 (*Von Hannover v. Germany*).

[54]See e.g., ECtHR, 11-10-1979, application no. 8348/78 & 8406/78 (*J. Glimmerveen and J. Hagenbeek v. the Netherlands*).

[55]See e.g., ECtHR, 25-06-2004, application no. 59320/00 (*Von Hannover v. Germany*); Hoge Raad, 18-01-2008, ECLI:NL:HR:2008:BB3210.

expression and information.

For such a balance of interests, I argue that it is vital to take the character and affordances of the Web into account. The Web may easily give rise to a form of 'overexpressiveness': as pointed out in section 4.3.3, the character and affordances of this medium easily give rise to more extreme and spur-of-the-moment expressions. As such, many users are likely to encode information about themselves, and others, that they would not entrust to many offline publications. Online, we see the publication of a wide variety of content, some of which is of such a nature that I think there remains little doubt that in these cases the interests of data subjects and/or their families should prevail over the freedom of expression and information. An example of such content that has been ordered to be erased by European courts, is a video of a disabled boy being bullied by teenagers.[56] Given the ease with which users encode information online, it is therefore important to consider whether the value of the content for the expresser and the public interest is proportional to the impact that the content has on the life of the data subject. Such a proportionality test is not something new or specific for art. 17 GDPR; throughout European case law we find examples of cases in which the freedom of expression and information is balanced against the (privacy) interests of individuals (see e.g., Von Hannover v. Germany and Axel Springer AG v. Germany[57]). However, I do argue that due to the combination of, on the one hand, the potential disinhibition effect that often underlies online expressions, and on the other hand, the potentially heavy impact of the Web as ubiquitous information source on the lives of individuals, this proportionality test should be given extra attention in the balance of interests regarding online information.

Additionally, the formulation of art. 17 (3)(a) is of interest for the balance of interests: "for *exercising* the right of freedom of expression and information"[my emphasis]. The gerund 'exercising' suggests a dynamic view on the freedom of expression and information, seeing it as an activity with potentially a beginning and an end. The phrasing seems to attribute a temporal scope to the act of expressing and thereby suggests that the information should be retained as long as it is part of an ongoing activity of expression, but may come up for erasure under art. 17 GDPR when the activity ended. An expression thus needs to be seen as having a sort of 'lifecycle'. It starts with the encoding phase, but does not end there: the idea of expressing oneself is that the expression reaches a certain audience. The possibility of reception is therefore an important aspect of the freedom of expression. The scope of 'exercising' the right to freedom of expression and information will thus need to be tied to the audience that the expresser intends to reach, both in scope (size of the audience) and in time (how long the audience should be receiving or being able to access the information). Moreover, separate from the expresser's intentions, the general public can also have a certain interest in exercising their freedom of information. When exactly

[56]Google blog about the case in the Turin court, `https://googleblog.blogspot.com/2010/02/serious-threat-to-web-in-italy.html`, last accessed 23-10-2018.

[57]ECtHR, 25-06-2004, application no. 59320/00 (*Von Hannover v. Germany*); ECtHR, 07-02-2012, application no. 3995/08 (*Axel Springer AG v. Germany*).

this sender-audience interaction as activity 'ends' and when the content moves into the 'inactive retention' phase is difficult to decide. Despite this difficulty, it is relevant to approach this as a process and take the impact of the passing of time into account.

Next to curbing indefinite retention, approaching the freedom of expression and information from a changing value over time has an extra advantage: when we take into account that our expressions may be longterm sticky notes to our informational persona, we may think twice before encoding information. While this in itself may not be such a bad thing, given for instance the implications of the disinhibition effect, it can also press too strongly on us. As I argued with Gorzeman in *Escaping the Panopticon Over Time*, the awareness of the fact that our online expressions may possibly be a lifelong representation of us, could cause a chilling effect due to which we refrain from encoding many of our views (Gorzeman & Korenhof, 2016). If we would live in a world of unlimited information retention, "the smartest way to survive is to be bland" (Ronson, 2016, p. 266). As such, an unlimited retention of information, could even eventually backfire on many of the values that are attributed to the freedom of expression, like raising diverse ideas and views, providing a view on the perceptions present in society, and/or giving individuals the freedom to ventilate their views and emotions (Gorzeman & Korenhof, 2016).

In sum, I argue that the relationship between the freedom of expression and information and art. 17 GDPR is nuanced. This is not to say that art. 17 GDPR is the perfect means to address problematic online expressions (I will touch upon some other or additional solutions in chapter 10), but it is not necessarily devastating to the freedom of expression and information either. The various elements raised in this subsection suggest a careful balancing of interests, where attention should be given to the proportionality and subsidiarity of a reference's continued presence, in view of an individual's interest in seeing it erased, as well as in the interest of the expresser and the general public in seeing it retained as it is. Yet, such a balance with the freedom of expression seems to be the main issue why art. 17 GDPR is so contested in the US. US authors believe the freedom of expression will too easily lose in the weighing of interests, while in the US the freedom to expression is seen as a fundamental value that has traditionally been granted a stronger protection than privacy (Bennett, 2012). On this level, the EU and the US maintain somewhat different perceptions on and appreciations of the right to privacy (while also having a certain overlap, see e.g., *The right to be forgotten: Reconciling EU and US perspectives* by Bennett (2012)). The differences in views between the US and the EU is not something that I can resolve in this particular study. Given the role of the GDPR, I will maintain an EU focus. I will delve deeper into the balance of interests with regard to the different technological applications in the next chapter — and hopefully convince some US authors of the possible merits of the approach that I suggest.

8.2.9.2 Exception (b)

"for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller"

Exception (b) concerns processing that is necessary for compliance with a legal obligation, a task carried out in the public interest, or when the controller acts on behalf of official authorities. In these cases, the information processing may continue despite a claim of the individual to a right to erasure. As such, exception (b) involves values and legal obligations external to the controller. Yet, in these cases it is still important that the controller critically evaluates to what extent the processing of the personal information is *necessary*; this ties in to the data minimisation principle of art. 5(1)(c) GDPR. Data minimisation entails a restriction to the collection of information to information that is "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed" (art. 5(1)(c) GDPR). The challenge of this exception lies in deciding to what extent the processing is really necessary. This will in many cases be subject to interpretation — and the controller and the individual invoking art. 17 GDPR will, in all likelihood, disagree on this.

In general, this exception prevents art. 17 GDPR from disrupting legal obligations that controllers may have. As such, this exception serves not only the controllers who have legal obligations, but also the public interest and the rule of law as it safeguards compliance with the law.

8.2.9.3 Exception (c)

"for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2) as well as Article 9(3)"

The baseline of exception (c) is that it safeguards the work of medical, social and health care services as well as health research. The limitation to art. 9(2)(h)-(i) and 9(3) GDPR confines the processing of personal information to those professionals that are bound by ethical standards or legal obligations to professional secrecy. Given the limitation to those cases of processing where the professional standards of secrecy apply, this exception is unlikely to play a role in cases that revolve around online publicly accessible and personally identifiable information. However, I think it is important to at least point out that public health is an important public interest that carries a significant weight in the balance of interests of art. 17 GDPR.

8.2.9.4 Exception (d)

"for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article

89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing"

Exception (d) reflects the importance that is attributed in the GDPR to the societal value of archiving information. While exception (d) protects the processing of information for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, this processing does need to have appropriate safeguards with regard to the processing of personal information (art. 89(1) GDPR). One of these safeguards is that the processing needs to be in accordance with the principle of data minimisation (art. 5(1)(c) GDPR). Safeguards may also entail the use of pseudonymisation and the like. If the appropriate safeguards are in place, the remaining content may fall under exception (d). The exception thus requires the controller to thoroughly consider the necessity of the processing of identifying personal information for archiving purposes. As such, this exception entails a balance of interests between the interests of the individual in the erasure of the information, and the societal value of the information that warrants its retention for the public good.

Given the importance of the retention of information for archiving purposes in the public interest, or for scientific or historical research purposes or statistical purposes, I take the interests protected in exception (d) to have a considerable weight in the balance of interests when weighed against the interests of the individual. Online, archives can be made accessible for the above mentioned purposes. However, a question here is whether the content should be publicly accessible online for everyone to fulfil its purpose. The processing purposes could, for instance, also be achieved by offering the information to researchers on request.

Personal information in publicly accessible online archives has already been the target of several court cases.[58] I will touch upon online archives and the balance of interests in section 9.3.1.2, especially in relation to views of courts on how this balance relates to personal information and the passing of time. However, delving in-depth into art. 17 GDPR's balance of interests in relation to specifically archives deserves a research on its own, and one that, in my opinion, requires an active involvement with researchers and archivists. For now, I suffice with concluding that art. 17 GDPR is written to respect research and archiving purposes, but in this, in combination with art. 89(1) GDPR, it latently embodies the view that archiving does not mean all-encompassing, indefinite and/or limitless information retention. Whether and to what degree art. 17 GDPR strikes a proper balance with this, I will leave to the scientific and archive community to form an opinion on.[59]

[58] See e.g., CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*); Cour de cassation de Belgique, 29-04-2016, C.15.0052.F/1.

[59] For example, the European Archive Group formulated such an opinion in its 'GUIDANCE ON DATA PROTECTION FOR ARCHIVE SERVICES - EAG guidelines on the implementation of the General Data Protection Regulation in the archive sector', `https://ec.europa.eu/info/files/guidance-data-protection-archive-services_en`, last accessed 03-09-2019.

8.2.9.5 Exception (e)

"for the establishment, exercise or defence of legal claims"

Exception (e) has some family resemblance to exception (b), as it supports the functioning of the law. This exception serves the legal certainty of parties who depend on the content for legal claims. Again, to what extent the (continued) processing of personal information is necessary, is a matter of interpretation and debate.

Exception (e) can have some relevance with regard to the type of content discussed in chapters 4 to 7. For example, in case of the Drunken Pirate, one can imagine that the school might lay a claim on MySpace to retain the picture for evidence purposes, if the data subject would sue the school for firing her. Another example could be the subway company in the Dog Poop Girl case, who may want to retain the video as evidence if they were to sue the Dog Poop Girl for the cleaning costs. While in such cases retention of the content can be important to establish, exercise or defend legal claims, it is doubtful that it is necessary to retain the content *publicly online as is*. Saving a screenshot of the content privately will likely suffice to prove and defend the legal claims. Such private storage of the content would itself not bring forth any of the identified problems. As such, exception (e) itself does not seem to be of major importance for the cases I examined. However, if the content retained under exception (e) would become part of a court case following a legal claim, this — if it spikes the interests of journalists, legal scholars, etc. — could lead to new online descendant objects that in turn may give rise to a new round of problems.

8.3 And we name it...

Naming is framing. This is the point where art. 17 GDPR takes an unfortunate turn. Starting out as "Right to be forgotten and to erasure"[60], its name was turned into "Right to erasure ('right to be forgotten')" in the final version of the GDPR. Yet, whatever the order, and despite the parentheses, the right is stuck with a double name and a double framing as a 'right to be forgotten' and a 'right to erasure'. With regard to this double naming, I follow partially in the footsteps of Jones and Ausloos who see art. 17 GDPR as the conflation of two different rights (Ambrose & Ausloos, 2013).[61] However, I differ from them with regard to

[60] See Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) /* COM/2012/011 final - 2012/0011 (COD) */, Document 52012PC0011, 2012, `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52012PC0011`, last accessed 05-08-2019.

[61] Although this analysis of Jones and Ausloos saw on an earlier version of art. 17 GDPR, I do not have the impression that this particular standpoint would be overthrown as a result of the changes in the final version, so I maintained their argumentation at the points where it still matched with the final version.

the meaning and the scope I attribute to this conflation. I will discuss the two names and frames, as well as their implications.

Right to be forgotten With the storage affordances of online information, it is not surprising that 'forgetting' was quickly taken to be a core problematic effect of the online tertiary memory (see e.g., Rosen, 2010). O' Callaghan and De Mars therefore argue that the 'right to be forgotten' is a useful metaphor, as it helps us to understand the problem of contemporary information technologies O'Callaghan & de Mars (2016, p. 263). Several scholars have underlined the importance of realising some form of 'forgetting' in the current IT era and have been looking into various ways and angles for the implementation of forgetfulness. Some of the notable examples of scholars who argued in favour of implementing a form of forgetting in ICT are Blanchette and Johnson, who delved into forgetting in relation to bankruptcy law, Dodge and Kitchen who worked on an ethics of forgetting to counter all too pervasive computing, and Mayer-Schönberger with his book *Delete: The Virtue of Forgetting in the Digital Age* (Blanchette & Johnson, 2002; Dodge & Kitchin, 2007; Mayer-Schönberger, 2009).

As right to 'be forgotten', art. 17 GDPR is often seen as related to the older French *droit à l'oublie* (Mantelero, 2013; Ambrose & Ausloos, 2013; Bartolini & Siry, 2016), the Italian *diritto all'oblio*, or the Spanish *derecho al olvido* (Bartolini & Siry, 2016). This right "has historically been applied in exceptional cases involving an individual who has served a criminal sentence and wishes to no longer be associated with the criminal actions" (Ambrose & Ausloos, 2013, p. 1-2). Forgetfulness, and in its footsteps *le droit à l'oublie*, fulfill a dual role here: they cover both forgetting by yourself and being forgotten by others (Ambrose & Ausloos, 2013, p. 14). In this conceptualisation, the right "is founded upon protections against harm to dignity, personality, reputation, and identity" (Ambrose & Ausloos, 2013, p. 14).

However, linking 'forgetting' to an individual right also led to critique on art. 17 GDPR that recurrently surfaced in the debate: calling art. 17 GDPR a right 'to be forgotten' can be seen as misleading or even plain wrong (see e.g., Koops, 2011; Powles & Floridi, 2014; Markou, 2015). 'Forgetting' is a concept that is generally understood as something that occurs in human memory. Removing a signifying object from the Web does not necessarily mean that people who have seen it forget it. Also, people can forget what they have seen even if the information is still online. As such, the erasure of online objects is something very different from human forgetting and the terminology is considered to be inappropriate for the online erasure of signifying objects. Moreover, instead of an internal natural form of forgetting, art. 17 GDPR entails a 'forgetting' that is artificially induced — even forced — by an external agent. Taken from this perspective, art. 17 GDPR is an interference with the autonomy of the information controller under the banner of a misplaced concept. Art. 17 GDPR has therefore often been described in the media as censorship or a means to rewrite history.[62]

[62]See e.g., Adam Thierer, "Europe's 'Right to Be Forgotten': Privacy as Internet

Some of this critique has likely hit a nerve with the EU legislator because in the final version of the GDPR, art. 17 was renamed to 'right to erasure', with the addition of '('right to be forgotten')' between parentheses and apostrophes. Yet, even between parentheses and apostrophes, 'right to be forgotten' remains part of the article's name and, in most of the literature and debate, is the term used to refer to the right rather than the final official main term 'right to erasure'.

Right to erasure This brings me to the other part of the right's name, 'right to erasure'. In its conceptualisation of the right to erasure, art. 17 GDPR is a mechanismal right (Ambrose & Ausloos, 2013, p. 14). In this form, art. 17 GDPR is closely linked to art. 12(b) DPD which should be seen as its predecessor (see section 8.2.6). With its focus on 'erasure', the right serves as an instrument to perform a concrete and material action aimed at a particular technological mediation. A concrete framing of art. 17 GDPR as right to erasure could therefore alleviate some of the criticism and problems that the right encounters in the public debate and in practice.

Double trouble Lastly, it is important to point out that the change in name of art. 17 GDPR not only changed the order between 'right to be forgotten' and 'right to erasure', but also the relation between the right and the names. The 2012 proposal for the GDPR had the form 'A and B', which can be read as one right (a right to A & B), but may even suggest two rights, a 'right to be forgotten and [a right] to erasure'. However, in the final version, the right's name took on the form 'B('A')', which suggest that B is an alternative name for A. Unfortunately, this synonym perspective seems to be at odds with the manner in which the GDPR itself presents art. 17 GDPR. Recital 66 contains the following phrase: "To strengthen the right to be forgotten in the online environment, the right to erasure should (...)". This formulation suggests that the right to erasure exists in support of the right to be forgotten. But if the right to erasure is formed by art. 17 GDPR, where then can we find this right to be forgotten? Or is the formulation merely poorly chosen and are the two names used interchangeably? Art. 17 GDPR thus suffers from ambiguity with regard to is name and frame: it is so far unclear whether the right merely has a double name, whether it is one right with a double function, or whether it represents two different rights that are collapsed in one article.

The double naming and framing has complicated much of the discussion about art. 17 GDPR, because the different character of the two frames remained present underneath the surface. Unfortunately, double naming and framing leads therefore

Censorship", *The Technology liberation front*, 2012. `https://techliberation.com/2012/01/23/europes-right-to-be-forgotten-privacy-as-internet-censorship/`, last accessed 24-10-2018; Jamie Grierson, "'Right to be forgotten' claimant wants to rewrite history, says Google", *The Guardian*, 2018. `https://www.theguardian.com/technology/2018/feb/27/right-to-be-forgotten-claimant-wants-to-rewrite-history-says-google`, last accessed 23-11-2018. Danny Sullivan, "Google Agrees To Complicated Worldwide 'Right To Be Forgotten' Censorship Plan" *Search Engine Land*, 2016. `https://searchengineland.com/google-to-censor-worldwide-sorta-243938`, last accessed 23-11-2018.

not to one stable entity, but to a Siamese twin that is fused together at certain points, while at other points it displays two different characters. Given the state of the right and the debate, I wonder if this right can survive as Siamese twin. The naming and framing of art. 17 GDPR as 'right to be forgotten' has taken the lead in media, academia and courts — while the right is being hinged on the mechanisms of art. 17 GDPR which are focused on erasure. Unfortunately, with the GDPR being effectuated as I wrote this dissertation, separating the twin so that one, or maybe even two, may live, may be too late.

However, there may still be a way out of this. This is the point where I diverge from Jones and Ausloos who argue that art. 17 GDPR is a conflation of a right to be forgotten based on the older 'right to oblivion' and a more mechanismal right to erasure. Instead, I attribute this conflation mainly to the right's name and frame and argue that it does not necessarily have to cover its functional mechanisms as a right. I will explain this further in the next section.

8.4 A right to ...

The rationale underlying the GDPR is that the "processing of personal information should be designed to serve mankind" (recital 4 GDPR). The protection of personal information therefore does not entail absolute rights, instead, "it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality" (recital 4 GDPR). In my opinion, this view is expressed in the mechanisms of art. 17 GDPR. While the right empowers individuals to battle certain extremes of the processing of their information, it also incorporates safeguards for the public interests, the interest of controllers, and the information processing interests of private individuals. However, even with these tentatively promising mechanisms, art. 17 GDPR comes with some challenges. The main problem of the right is that it is still unclear what the right is, what should it resolve, and how. In the debate surrounding the right to be forgotten, we therefore see different views surfacing that place the emphasis on different elements (erasure, forgetting, time, etc.). The uncertainty in the debate of what the right is, and it meant to do, seems to be at least partially caused by the right's double naming and often framing.

While 'forgetting' in relation to art. 17 GDPR can serve as a useful concept (as I have argued in my essay *Forgetting Bits and Pieces: An Exploration of the 'Right to Be Forgotten' as Implementation of 'Forgetting' in Online Memory Processes* (Korenhof, 2013)), the conceptualisation of art. 17 GDPR as a 'right to be forgotten' overall seems to steer our view in suboptimal directions. First and foremost, focusing on art. 17 GDPR as 'right to be forgotten' places the emphasis on issues with regard to the 'everlasting' memory of the Web. As I have shown in the previous chapters, the long retention time of online signifying objects is not the only issue that it raises. A focus on 'forgetting' may thus easily cloud a significant part of the problems and their main contributing factors, such as space and proportionality. This, while, in its guise as 'right to erasure', art. 17

GDPR may be able to address these issues. Additionally, the conceptualisation of art. 17 GDPR as 'right to be forgotten' is highly metaphorical. While metaphors can certainly help to clarify issues, in the case of art. 17 GDPR, the 'forgetting' metaphor seems more likely to be obfuscating the issues, as well as the debate surrounding the right. I therefore argue that it is better to take the mechanisms of the right as a point of departure for our further investigation. The mechanisms of the article express a particular functionality: they are focused on giving the individual a certain degree of control over her personal information by means of erasure. However, even 'erasure' in the context of art. 17 GDPR seems to have a somewhat metaphorical character: 'erasure' can take on diverse forms, some of which are technically not even forms of erasure (see section 8.2.6). While this is likely not beneficial for the clarity of the right, I argue that the vague concept of erasure may be a blessing in disguise because it allows us to explore a wide range of removal options to resolve the issues at hand.

By requesting the erasure of particular content, which can range from a single signifying object to a huge dataset, the individual can try to (re)claim control over her personal information. She can do this when she believes that the further processing of her personal information is no longer necessary for the controller (ground (a)), when she revokes her consent to the processing (ground (b)), when she objects to the processing based on her particular situation (ground (c)), when the content is unlawfully processed (ground (d)), when it needs to be erased in accordance with the laws of a Member State (ground (e)), or when the content was shared by her when she was a child in relation to the offer of information society services (ground (f)). With these grounds, the data subject is given a relatively strong retroactive discretionary power over the processing of information that relates to her. However, the interests of the subject do need to be carefully balanced against the interest of the controller and the general public in the ongoing processing of the information. As argued in section 8.2.9, many of the exceptions to art. 17 GDPR give rise to a sliding scale with a tipping point somewhere along the balance of interests. Combined with the various forms of erasure that art. 17 GDPR allows, the right offers a lot of room to manoeuvrer in order to resolve the identified problems in a manner that can do justice to the various interests involved.

Given the above, I suggest focusing from this point onwards on art. 17 GDPR as a 'right to erasure'. From this perspective, art. 17 GDPR constitutes a certain instrument that the individual can use under specific conditions to shape her informational persona by means of erasure requests. In the following chapter, I will continue to approach art. 17 GDPR from its functionality, erasure, and deepen its relation to the technological mediation of personal information.

Chapter 9

Art. 17 GDPR and the problem narrative

Contents

9.1 Introduction

A rose by any other name would smell as sweet

William Shakespeare, *Romeo and Juliet*, 1595

A rose has certain qualities that define it as a substance, whatever name we give it. Shakespeare had a point there. In the previous chapter, I therefore have been tracing the qualities of art. 17 GDPR to figure out what the article is and does. Yet, like a rose, art. 17 GDPR looks nice at first glance, but pricks when you try to touch it; the article comes with a handful of questions and complications — not in the least, its name. Given the fact that art. 17 GDPR has already come into force and will be used for years to come, it is important to conceptualise the right in the best possible way to apply it in an effective and just manner. At the end of chapter 8, I therefore proposed to move beyond the unfortunate double naming of art. 17 GDPR and instead focus on the right's functionality (erasure) and deepen its relation to the technological mediation of personal information. In this chapter, I will take up this challenge and explore how we should understand and apply art. 17 GDPR in order to successfully address the problems identified in chapters 4 to 7.

With the character of the problems and the core mechanisms of art. 17 GDPR clarified, I think it is important to delve towards a deeper understanding of the problems in relation to the character of the right. This will allow us a clearer view on how and when to apply the right, as well as how to balance the interests of the stakeholders. Please note that for the purpose of this study, I will maintain a general focus on art. 17 GDPR's capability to address the problems by allowing data subjects to request the erasure of particular information relating to them. The assessment of art. 17 GDPR in this chapter therefore entails only a partial analysis with its focal point on the data subject realising erasure. The full analysis is a dissertation on its own, and one that has already been eloquently written by Ausloos (2018).

I will assess the capability of art. 17 GDPR to address the problems identified in chapters 4 to 7 by combining the many and intertwined issues that I discussed into a bigger picture that connects to some of the main notions underlying the right. This bigger picture will then serve as a frame to evaluate in more detail the problem-solving potential of art. 17 GDPR, and especially with regard to the balance of interests that goes with it. In turn, the analysis of the previous chapters provides me with guidance on when and how to plausibly apply art. 17 GDPR in order to resolve or reduce these issues. I will examine the functionality of the right to erasure in addressing issues in the three main technological applications and the phenomenon that I examined: (1) web pages (chapter 4), (2) social media applications (chapter 5), (3) search engines (chapter 6) and (4) virality (chapter 7). Lastly, I will conclude this chapter with an overall evaluation of the problem-

solving potential of art. 17 GDPR. First, however, I will focus on the bigger picture.

9.2 The problem narrative

In this section, I delve more deeply into what art. 17 GDPR is, or ideally should be, given what we know of the problems that it needs to address. I will start by discussing what kind of right art. 17 GDPR (ideally) is. I will explain that the most promising approaches to art. 17 GDPR all relate to autonomy and identity. For reasons that I clarify in the first subsection, I continue by discussing the main elements of Ricoeur's identity theory. I will elaborate on how this connects to the theory of technological mediation and the identified problems. Based on the previous, I will suggest a particular conceptualisation of art. 17 GDPR. This then will serve as a foundation for the following sections, where I will investigate how to apply art. 17 GDPR to the problems identified in the previous chapters and assess the right's problem-solving potential.

9.2.1 Type of right

In the many discussions I got involved in during the writing of this dissertation, I encountered scholars who argue in favour of understanding art. 17 GDPR as a right to mitigate only economic, or at least clearly financially quantifiable, risks and harms. Taking in a financial harm perspective to understand and apply art. 17 GDPR has a certain advantage: it makes the balance of interests relatively clear as the erasure should only be applied if such a harm can be demonstrated. For instance, take the case where the online presence of a particular piece of personal information consistently frustrates an individual's chance to get a job. The harm is clear here: it is an economic harm for the individual, but also possibly for the state as the individual may need to depend on social welfare. The extent of the harm can relatively easily be calculated in financial terms. The financial harm perspective allows us to evaluate rather tangible consequences against each other when balancing the interests of the different parties that have a stake in the matter. Despite this advantage of the financial harm perspective, the mitigation of such harms does not seem to be the target of art. 17 GDPR, nor to be the best perspective for understanding and applying art. 17 GDPR given the identified problems. EU Justice Commissioner Reding argued that the introduction of the GDPR and more specifically art. 17 GDPR, is about giving individuals more *control* over their personal information[1], and specifically in relation to their own *identity*[2]. Moreover, if we can take the *Google Spain* case as a precedent for the

[1]European Commission press release, Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses, `http://europa.eu/rapid/press-release_IP-12-46_en.htm`, last accessed 19-07-2018.

[2]Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*`http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

application of art. 17 GDPR, we see in recital 96 of the verdict that the CJEU confirms the view that suffering a concrete harm is *not a necessary requirement* for the erasure of information. Here, the CJEU argues that any signifying object may be up for erasure, even if it does not actually give rise to any prejudices with regard to the data subject.[3] This is consistent with how I read the text of art. 17 GDPR. As I concluded in chapter 8, the article gives the individual a certain degree of control over what information others process about her.

However, even with this in hand, it is not necessarily clear what kind of right art. 17 GDPR is, and how we should understand its goals. In DPD and GDPR literature, we can find three main perspectives that can be used to make sense of art. 17 GDPR: a privacy perspective, a data protection perspective, and a identity perspective. In order to explain these perspectives, I connect to the work of diverse scholars who wrote about privacy, data protection and/or identity rights, albeit it not necessarily about art. 17 GDPR. In the work of some authors, these perspectives overlap partially or even fully[4], while others aim to make a strict distinction between (two of) the perspectives[5]. Discussing the various views on these perspectives in detail and resolving their differences lies outside the scope of this study. What I will do, is concisely discuss what these three angles can tell us about what art. 17 GDPR should protect.

9.2.1.1 A privacy right

Starting with the privacy angle. This is the point where I cannot avoid the concept of privacy any longer. Those readers experienced in reading reflective literature on the GDPR, or more particularly art. 17 GDPR, may have noticed that this study so far hardly touched upon the one particular concept that plays a dominant role in the majority of the debate: *privacy*. I have two reasons for not touching upon the notion of privacy earlier. The first is that I wanted to approach the problem spectrum that art. 17 GDPR deals with as openly as possible. Framing it from the start in the context of privacy could obscure some of the mechanisms and problems, or even colour them unjustly. Secondly, 'privacy' is a complex notion and has been the topic of many books and research papers (see e.g., Westin, 1970; Schoeman, 1984; Solove, 2005; Hildebrandt, 2006; Roessler, 2005; Koops *et al.*, 2017) — a diversity and complexity that I cannot fully and properly address in this study. Privacy covers different fields and is not reducible to personal information; it also entails spatial privacy, bodily privacy, decisional privacy, and relational privacy (Roessler, 2005; Rouvroy & Poullet, 2009). Even what privacy is, a concept, a value, a phenomenon, or a right, is a topic of discussion (Hildebrandt, 2006).

[3] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §96.

[4] For example, Agre as well as Hildebrandt argue that privacy is about the process of identity building (Agre, 1997; Hildebrandt, 2006).

[5] For example, de Andrade stresses that privacy and identity entail different rights (de Andrade, 2014, p. 67). Yet, he holds a relatively narrow definition of privacy as he approaches the right to privacy mainly "from a more classical and delimited perspective as a right to opacity and seclusion" (de Andrade, 2014, p. 67).

The different views on what privacy is, show a wide array of approaches and applications. These do not necessarily conflict, but can coexist as different cases of privacy. Hildebrandt therefore argues that these different cases and contexts of privacy could best be seen as having a 'family resemblance' to each other in Wittgenstein's sense, without having a overarching definition that covers them all (Hildebrandt, 2006). I agree with Hildebrandt and follow her in her argumentation that focusing on a strict definition of privacy would reduce its value for practical use. Views of what should be private and what not, differ per culture and tend to change over time. An all-too-strict definition of the concept may render it useless when changes occur in our culture, technologies and/or lifestyles.

While privacy is an underdetermined concept, there is some agreement on its virtue and goals. Scholars like Westin, Agre, Roessler, Hildebrandt, Rouvroy and Poullet all value privacy as a means to safeguard the autonomy of the self (Westin, 1970; Agre, 1997; Roessler, 2005; Hildebrandt, 2006; Rouvroy & Poullet, 2009). In this, privacy has an instrumental value (Westin, 1970; Rouvroy & Poullet, 2009; Porcedda, 2017). It preserves individual autonomy by establishing a certain margin of freedom in societal life. In some cases, the concept of privacy is directly intertwined with identity, as for instance is done by Agre, who defines the right to privacy as "the freedom from unreasonable constraints on the construction of one's own identity" (Agre, 1997, p. 7). While privacy is generally defined in a manner that concerns safeguarding the autonomy of the individual, it is important to note that it benefits not only the individual: the development of individuals and their views is of vital importance for a healthy functioning of society. Privacy serves the public interest by providing individuals with the freedom and space to develop their own perspectives and critical opinions which are important for the sustainability of democratic society (Westin, 1970, p. 24). As art. 17 GDPR aims to provide people with a certain degree of autonomy with regard to their personal information, the right seems to at least have something of a privacy right. In particular, the right's user centred approach is especially valuable for the safeguarding of individual autonomy (Weber, 2011, p. 125). Also, what is notable, is that art. 17 GDPR's functionality shows similarities to Westin's definition of privacy: Westin defines privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" (Westin, 1970, p. 7). However, this also brings in the next perspective: data protection.

9.2.1.2 A data protection right

Westin's definition of privacy highly influenced the development of data protection law (Porcedda, 2017). While this has strong ties to privacy, several authors argue that the right to data protection should be seen as a separate right (see e.g., Rodotà, 2009; Porcedda, 2017). What supports this view is that in the EU Charter of Fundamental Rights, the right to data protection is codified as an autonomous right in art. 8. With this, the right to data protection is separated from the right to respect for private and family life which is codified in art. 7. Porcedda argues

that "both rights help the person keeping (solid) control of the process overseeing the creation and maintenance of her or his identity and, relatedly, dignity and autonomy, but each does so in a different manner and expresses different core areas" (Porcedda, 2018, p. 295). The right to data protection is taken to have a specific focus on addressing issues that come with the development of new information technologies (Rodotà, 2009; Porcedda, 2017). According to Porcedda, one of the main values that underlies the right to data protection, is to have "control over the portrayal of one's identity (and related personality) to society, and the consequences that can ensue from such portrayal" (Porcedda, 2018, p. 300-301). This shows a strong overlap with art. 17 GDPR as an identity right, which I will discuss in the next subsection. Viewing art. 17 GDPR as a data protection right that establishes a certain degree of control over the portrayal of one's identity in the light of new information technologies ties in to the mechanics of art. 17 GDPR, as well as to the issues identified in chapters 4 to 7. I therefore take the data protection perspective as a relevant perspective for understanding art. 17 GDPR. Additionally, and more importantly, given that art. 17 GDPR is part of the General *Data Protection* Regulation, I conclude that it is at least meant by the legislator to be a data protection right and has its focus on protecting individuals against certain impacts of information technology.

9.2.1.3 An identity right

Lastly, we can also view art. 17 GDPR specifically as an identity right. De Andrade describes identity rights as the right to have "attributes or the facets of personality which are characteristic of, or unique to a particular person (such as appearance, name, character, voice, life history, etc.) recognized and respected by others" (de Andrade, 2014, p. 67). While sharing some similarities with the privacy and data protection angles, there is a relevant difference between Agre's definition of the right to privacy (see above) and his definition of the right to identity: the first is focused on a freedom from constraint (others should not unnecessarily restrict the data subject), while he attributes the right to identity a focus on affirmation (the recognition and respect of others). However, later in de Andrade's text this respect and recognition of others seems to have taken shape in a right to identity that entails a right to be oneself where others do not use a data subject's personal attributes "without authorisation in ways which cannot be reconciled with the identity (...) that they wish to convey" (de Andrade, 2014, p. 68). In this form, de Andrade's right to identity seems conceptually closer to his definition of the right to privacy. I will not delve deeper into this, but will leave it in the open with the remark that the demarcation and overlap between privacy and identity rights depends on the definition of both types of rights.

Leaving the demarcations between the different types of rights aside and focusing on art. 17 GDPR specifically from an identity perspective, de Andrade offers an interesting conceptualisation for art. 17 GDPR as an identity right. He argues that the right ideally should help individuals to develop themselves over time without having to fear systematic stigmatisation in the here and now

related to their past actions and opinions (de Andrade, 2014). It should be, as de Andrade phrases it, a "right to be different from oneself, namely one's past self" (de Andrade, 2014, p. 69). While de Andrade offers a very interesting and valuable view, it is focused on one particular angle: change of the individual over time. This does not account for a significant part of the problems that I identified in chapters 4 to 7; the problems are more diverse than information that is outdated or decontextualised because the individual herself changes. In the technological mediation, online content can be resemiotised, disproportionally represented, overrevealed, etc. While the 'right to be different from oneself'-perspective surely touches the heart of a part of the issues, I argue that given the diversity of the problems and the role of technology herein, it seems to still fall a bit short of fully reflecting and addressing the issues at hand.

9.2.1.4 A family resemblance

While it seems relatively clear that art. 17 GDPR falls at least under the data protection laws, it also has particular characteristics that strongly relate to privacy as well as identity rights. It is difficult to brand art. 17 GDPR specifically as one of these three. What makes this more difficult is that, depending on the definitions used, the distinction between privacy, data protection, and identity rights in many cases remains hazy. Referring back to Hildebrandt's view that the different cases and contexts of privacy share a certain family resemblance, I think it is fair to conclude that these three rights share a certain family resemblance: all three perspectives relate to an individual's autonomy with regard to her own identity. Moreover, the motivation of the EU legislator to introduce art. 17 GDPR connects to the subject's control over her own identity as well. I therefore propose to not stop here but to further explore art. 17 GDPR in relation specifically to identity and autonomy in combination with the full scale of the problems. Instead of working further top-down by framing art. 17 GDPR as either a privacy, a data protection, or an identity right and applying it from that perspective to the problems, I suggest picking up the notion of identity and then working with this notion from the bottom-up, from our understanding of the problems, towards a conceptualisation of a right to erasure that will provide us with a stronger grip on when and how to apply it.

In order to show how a conceptualisation of art. 17 GDPR in terms of identity construction touches the identified problems, it will be helpful to further elaborate on a framework of identity construction that can be applied to the four cases of technological applications. With its main functionality as an identity-related right, I suggest building further upon the work of the philosopher Ricoeur, who aimed to construct a theory of personal identity that allows us to understand our self in relation to our existence over time and to the surrounding world. Ricoeur's view of personal identity encompasses sameness as well as change of the individual (Ricoeur, 1992, p. 113). His bridge between sameness and change lies in the narrative as an overarching story of our lives. While, again, time receives particular attention in relation to identity, Ricoeur's theory is rich and

has sufficient potential ties to technological mediation for me research the issues beyond the solely temporal effects. As Ricoeur combines the phenomenological and the hermeneutic in his narrative theory, his views on identity as a narrative construction allow me to delve more deeply into the character of the problems. Although it is in itself not new to use the concept of narrative identity to discuss the implications of information technology for individuals (see e.g., de Mul, 2002; Hildebrandt, 2006; de Andrade, 2014; Buitelaar, 2014) and the potential use of specifically art. 17 GDPR to address these issues (see e.g., Burkell, 2016; Tirosh, 2017)[6], it is worthwhile to revisit this approach for the evaluation of art. 17 GDPR, but now with an explicit focus on technological mediation and its hybrid intentionality.

De Mul already made some valuable first steps in applying Ricoeur's concept of narrative identity to the Web in his *Cyberspace Odyssey*, first published in 2002.[7] While Ricoeur mainly focuses on narrative identity in the form of a single story, de Mul argues that, especially in the online world, more fluid and interactive narrative structures take shape (de Mul, 2002, p. 209-214). I will pick up where he left off. I modify, deepen and extend his Odyssey so that it can help us with the application and the balance of interests of art. 17 GDPR. I will do this, on the one hand, by applying it to the current Web (de Mul mainly discusses personal home pages, and does not delve into new applications). On the other hand, I will bend the concept of narrative identity further towards the theory of technological mediation and apply it to the technological applications discussed in chapters 4 to 7. The analyses of the impact of the technological mediation on the informational persona as set out in this study serve as a base upon which I 'mount' the concept of narrative identity and assess the implications of the mediation for the construction of a materialised narrative identity.

The reason for continuing this journey with Ricoeur is that, next to the explanatory merits of his theory, it also at crucial points ties in with the view on technology that forms the backbone of the present study. What is characteristic for Ricoeur, is that he argues that self-knowledge can only come through our understanding of our relation with the world we live in and in our interactions with others; we understand ourselves through our experiences by means of our reflexive consciousness (Pellauer & Dauenhauer, 2016). This view on our self-understanding has a crucial similarity to the view on technology presented in chapter 3, especially with regard to Stiegler's theory of technology as tertiary

[6]It is important to remark here, that Tirosh points out a critical characteristic of art. 17 GDPR, one that he labels as a flaw. Art. 17 GDPR is focused solely on individuals and as such it does not address information relating to groups and collectives. This characteristic is not addressed here, because the scope of this study is limited to individuals. However, this focus of art. 17 GDPR on individuals is an important restriction, and its implications are important to consider in further research. I will briefly touch upon this in chapter 10.

[7]De Mul also rightly points out some issues with Ricoeur's narrative model. However, because I base myself loosely on Ricoeur, this is not the place and time to get into a discussion on the scope and limits of Ricoeur's view. I therefore would like to refer readers to de Mul's *Cyberspace Odyssey* for further reading (2002). There is an English translation available of this book, published in 2010. Unfortunately, I could not get a hold of this version.

memory; the technological world around us forms the basis on which we understand our relation to the world, each other, and ourselves. Where Stiegler takes the world around us as an inevitable protention for our memory, Ricoeur finds that there is no unmediated self-understanding. While Ricoeur does not explicitly focus on technology in the direct sense of the word, he does focus on the impact of language and the necessity of interpretation for our self-understanding. He pays particular attention to the manner in which signs stored in memory and literary traditions affect our self-understanding (Pellauer & Dauenhauer, 2016). Moreover, Ricoeur argues that with the encoding of information, the text itself becomes the object that transfers meaning, thereby sidetracking the intention of the speaker (Pellauer & Dauenhauer, 2016). This connects to Verbeek's view on the hybrid intentionality of technology and its users (I will strengthen this bridge between specifically the technological intentionality and Ricoeur's view on identity in subsection 9.2.3). Before we move to the relation between technological mediation and narrative identity, I will first explain the basics of Ricoeur's theory of identity in the following subsection.

9.2.2 Narrative identity

Identity research brought forth the idea that the human self is constructed in terms of a narrative, a coherent story that individuals tell about their lives (see e.g., Mink, 1978; King, 2000; McAdams *et al.*, 2006). However, due to the aforementioned reason that the work of Ricoeur connects at crucial points to the view on technology that forms the framework of my research, I shall focus specifically on his view of the narrative identity.

Ricoeur's theory of personal identity leans on two pillars: on the one hand it leans on *idem* which means sameness, and on the other hand on *ipse*, which refers to the self or selfhood. Our *idem*-identity consists of a set of significations that establish a particular identity (Ricoeur, 1992, p. 2). It implies a sameness of our identity: we can be recognised as 'the same' over time, or 'the same' as others (de Vries, 2010, p. 74). It is a "*what* I am" (de Vries, 2010, p. 74). While we pass through time and space, we remain the same entity or can be recognised as being the same as others. For example, like other people from the Netherlands, I am recognised as 'Dutch' (a sameness to other people). Also, despite the fact that I age, I am recognized as 'Paulan' (a sameness over time). Even when a great distance in time challenges our resemblance to our earlier selves, this change is bridged by the "*uninterrupted continuity* between the first and the last stage in the development of what we consider to be the same individual" [emphasis in original](Ricoeur, 1992, p. 117). There is thus a certain *permanence in time* of identity (Ricoeur, 1992, p. 117). This permanence is strongly expressed in (but not exclusive to) our *idem*-identity (Ricoeur, 1992, p. 116).

Contrary to *idem*, our *ipse* identity "implies no assertion concerning some unchanging core of the personality" (Ricoeur, 1992, p. 2). The *ipse* is shaped by our reflective consciousness; we shape our *selves* by recognition over time and realising options of several possibilities. It is a "*who* I am" (de Vries, 2010, p. 74).

As such, the personal identity is never a full *idem* like an unchanging object is, but always also an *ipse*, a selfhood that is actively pursued by the individual who makes choices, expresses herself and chooses certain actions and events over others, changing over time. In this, the individual will also often identify herself with ideals, values, norms, models, others and (sub)cultures. These are the 'acquired identifications' (Ricoeur, 1992, p. 121). Because the individual recognises herself in these elements, she will internalise them. In turn, she will express them and thereby these elements will serve as signals for recognition (Ricoeur, 1992, p. 121). Others use these signals to form a view of the individual's identity and react to her accordingly (Goffman, 1959).

By recognising herself as being a particular individual or belonging to a certain group, the individual actualises certain possibilities and thereby shapes her identity. These actualised possibilities give rise to her character. The character is the "set of lasting dispositions by which a person is recognized" (Ricoeur, 1992, p. 122). The character consists of a set of distinctive signs with which the *ipse* announces itself as *idem* (Ricoeur, 1992, p. 121). If our character traits and habits are persistent, we can see a transformation of the *ipse*, the self, into *idem*, sameness (de Mul, 2002, p. 204). The permanence of character entails a full overlapping of the *ipse* with the *idem*. Ricoeur therefore finds the core of our selves in the dialectical relation between the *idem*-identity, sameness, and the *ipseity*, the selfhood. This is where the narrative comes in. The narrative mediates the constitution of our personal identity between "the pole of the character, where *idem* and *ipse* tend to coincide, and the pole of self-maintenance, where selfhood frees itself from sameness" [emphasis in original] (Ricoeur, 1992, p. 119).

The narrative identity oscillates between the poles of *idem* and *ipse* and connects events by means of *emplotment*. Emplotment is the ascribing of a plot to a set of separate events. Drawing on Aristotle, we can say that a plot is "the arrangement of the incidents" (Butcher, 1951, p. 25). Aristotle refers to the plot as the imitation of action and of life, because life itself "consists in action" (Butcher, 1951, p. 27). Emplotment is thus a 'configurational act' that mediates between the actual events and the narration of these events, by organising these events in a particular manner (Carr, 1991, p. 64). As such, emplotment "allows us to integrate with permanence in time what seems to be its contrary in the domain of sameness-identity, namely diversity, variability, discontinuity, and instability" (Ricoeur, 1992, p. 140). The plot entails the mediation between permanence and change (Ricoeur, 1991, p. 77). Ricoeur states: "the narrative constructs the durable character of an individual, which one can call his or her narrative identity, in constructing the sort of dynamic identity proper to the plot [*l'intrigue*] which creates the identity of the protagonist in the story" (Ricoeur, 1991, p. 77). The character itself is therefore a plot (Ricoeur, 1992, p. 143). The narrative identity is always a construction because the configuration of the narrative brings heterogeneous elements together in a whole (Ricoeur, 1992, p. 141). It brings structure and coherence to our identity. Adding something to the narrative signals significance. The narrative is therefore never neutral (Ricoeur, 1992, p. 115).

We do not only shape our narrative identity, but our narrative configuration

is also reflexively applied to the self (de Mul, 2002, p. 206). Ricoeur argues that self-knowledge entails only an indirect knowledge, "through the detour of cultural signs of all sorts, which articulate the self in symbolic mediations" (Ricoeur, 1991, p. 79). Knowledge of the self is an interpretation in which the self "figures itself as this or that" (Ricoeur, 1991, p. 79). The narrative thus plays a mediating role in the way in which we experience and construct our identity. However, our role as author is limited, because we cannot control the events that happen in our lives. At most, we can take up the role of coauthor and author the meaning given to these events in our narrative identity (Ricoeur, 1992, p. 162).

Lastly, it is important to point out that the individual is never a 'lonely isle'; our identity always takes shape in relation to others. We attach and detach ourselves to others, we embrace, accept, or resist social categories, we learn from others or position ourselves against their views. Even the hermit is defined by her relation to others, namely by her relation of detachment. The shaping of our *ipse* thus always takes place in an interplay of acceptance of and resistance to social attachments with others. Our narrative identities are intertwined with those of others, because we often play a role in their narratives, while they play a role in ours. The self is therefore part of a weave of personal narratives (de Mul, 2002, p. 208).

9.2.3 The technical storyteller

Our narrative identity helps us to make sense of our lives and identity. The plot of the narrative identity tells what is important about you, and what makes you *you*. While narrative identity is a concept that is often used in relation to internal processes and on an autobiographical level, Ricoeur already brought material inscription as well as other people into the construction of the narrative identity. Following this, de Mul, Hildebrandt as well as de Vries explored the use of Ricoeur's theory in relation to identity in digital information technology (de Mul, 2002; Hildebrandt, 2006, 2009; de Vries, 2010). Another view is added by Coeckelbergh and Rijers, who explore the narrative capacity of technology itself (Coeckelbergh & Reijers, 2016). Drawing inspiration from this combination of approaches and tying in to the view on the presence of information as put forward in the present study, I would like to take Ricoeur's concepts further specifically in the direction of technological mediation and its implications for the presentation of a view on the informational persona. The view on our personal information given to audiences (in the case of this study, Web users) can tell people a particular story about who we are and thereby affect how others understand us. Also, it can affect how we see ourselves by, for instance, reminding us of certain things or by confronting us with a particular view on ourselves. As others as well as ourselves engage with this materialised story about us, it is essential to examine how this presented image comes to correspond to or diverge from what we ourselves take to be our story. I therefore propose to use Ricoeur's concept of narrative identity, and its relation to sameness and change of an individual, in the context of the manner in which the tertiary memory establishes a particular presence of the informational persona. Employing Ricoeur's conceptual 'toolkit' in a material

context allows me to look more closely into the potential problematic differences between the narrative's plot as presented by the mediating technology and the referent's own sense of her narrative identity (which, in turn, is necessarily affected by the protentive workings of the narratives materialised in the tertiary memory, a relation examined by de Vries (2010)).

When we transfer personal information to a mediating technology, the information we materialise gains a certain autonomous existence, separate from its author. The references transcribed into these signifying objects make up the 'what' of the informational persona: e.g., the referent is German, female, a truck driver, a vegetarian, a Muslim, a violist, a gamer, etc. However, I have shown in chapters 4 to 7 that these objects can imbue certain references with a stronger or weaker presence, combine references, and/or affect their meaning. While the references in the signifying objects represent a 'what' of the informational persona, the manner in which they are presented by the signifying object, and in turn its context and arrangement between other objects, gives shape to a certain view on the informational persona. With this, technology 'tells' us more than solely the content of the object. By increasing, decreasing and connecting certain references, the mediating technology weaves the 'what' into a narrative that suggests a certain sameness and change of the referent. It thereby gives rise to the materialisation of a 'who' of the informational persona. As the arrangement and context of the signifying objects affect the story that is told about the identity of the referent, I suggest approaching the interplay of the presence and context of signifying objects as a shaping of 'the narrative identity' of the informational persona — or in other words, a materialised narrative identity of the referent. By transferring some of its affordances and characteristics to the personal information that it holds, the mediating technology thus not only presents a particular *materialised* version of the narrative identity, but also co-shapes the construction of the narrative that it helps to bring forth. With this, the mediating technology impresses some of its directionality, and thereby its intentionality, on the construction of the materialised narrative identity. While technology does not construct a narrative in the classical sense, I therefore argue that it does take part *in* the construction of the materialised narrative identity and thereby affects the story that is told. This material narrative may diverge from the referent's own sense of her narrative identity. I will explain this by means of an example.

In the case of an autobiographical book, the initial narrative is materialised on paper. Due to this materialisation, the narrative can reach audiences that do not share the same space-time zone as the original human narrator. As such, I argue that the book becomes a placeholder narrator. In this role, the book affects the audiences of the narrative by on the one hand stretching the narrative in space and time, while on the other hand restricting the narrative's audience to members who can read. However, it does more to the narrative than affect the narrative's potential audience. Once printed on the paper, the book transfers its relatively static materiality to the narrative. Written down, the narrative remains as it is. Changing the narrative, even minor details, would leave substantial marks on the paper, that, in turn, remain part of the narrative. While the content in the book

may narrate about sameness and change of a particular referent, on an overarching level, the book presents *this* particular story about sameness and change. Changes in the individual's own narrative identity as a result of new events in her life and her self-reflections thereupon, will not be part of the plot presented by this mediator. The book's material directionality thus impresses an overarching inclination to sameness on the presented identity of the referent.

Moreover, the book presents the narrative in a certain wrapper, often with additional text on the back-flap etc. This wrapper affects how the audience approaches the narrative and is likely to frame it. It even affects whether we pick up the book in a bookstore in the first place. By presenting the content in a particular context and manner, the mediating technology can thus affect the meaning of the narrative. Additionally, the paper materialisation of the narrative affects the interaction between the audience and the narrative. With the narrative materialised on paper, the book allows the audience to navigate its plot as it likes, to shift through the narrative at their own pace, to read the end first, etc. As such, the mediating book allows a certain flexibility to the order of the narrative that is not the result of an intentional activity of the human author to establish her identity. Instead, it is the mediating technology that impresses its own intentionality on the narrative by presenting it in a certain manner and imbuing it with particular affordances. Such an externalised narrative can be problematic for the referent because it may attribute a certain quality to the self that the individual may not consider to be representative for her, or at least less than this narrative suggests.

Given the impact of technological intentionality on the narrative act and its plotlines, I argue that we can understand the mediating technology in the role of placeholder narrator as a second layer narrator, narrator_2. Going back to the example of the book, let us say the book, in turn, is placed in a library, in the history department. As such, the building and the grouping of the library tell us at least two things; one, the autobiography is considered to be of interest to the general public, and two, the book is considered to be of historical interest. With this, the library wraps the content in another narrative layer by affecting the relation between the content and the audience. The library is narrator_3. And so on.

This layering of narrators entails a complex hybrid intentionality in which the impact of the human and the technological narrator is intertwined in different manners, with various degrees of human and technological intentionality. For instance, narrator_1, the human narrator, can deliberately employ a technological narrator as an instrument for her narration in order to remember a certain event or to reach a particular audience. However, if she fails to prevent, or sufficiently take into account, an unwanted impact of a placeholder narrator, something that happens quite easily as we have seen throughout this study, the mediating technology impresses a relatively strong technological intentionality on the narrative which is unintended by the author.

9.2.4 Reconfiguring the technologically mediated narrative

As discussed in the previous subsection, the technology mediating the materialised narrative identity imprints its intentionality on the narrative act, thereby affecting the audiences, the narrating voice, the time and the place of the narrating, as well as the plot of the narrative itself. The points on which a mediating technology can affect the narrative and undercut an individual's control over her materialised narrative identity, tie in to Westin's definition of privacy which focuses on the control that people have over the manner in which information about them is revealed to others. In this, the control over personal information is not just about the *who* with whom information is shared, but it is also about the *when*, the *how*, and *the extent* to which this information is shared. The mediating technology affects the sharing of information on all these points and can lead to unwanted and/or unforeseen consequences for the narrative identity of subjects as it is given shape in the outside world.

Ideally, art. 17 GDPR, can be of help by reconfiguring this externalised narrative. In this light, it is helpful to consider Hildebrandt's view on privacy rights. Hildebrandt argues that the core of what these rights should protect, is an 'indeterminacy of self-identity' (Hildebrandt, 2006). If we look at art. 17 GDPR from this perspective, we can say that art. 17 GDPR should protect an individual's freedom to construct her own narrative identity against a determinacy of her identity raised by the processing of personal information. This is where I argue that an important role for art. 17 GDPR should lie. Art. 17 GDPR should help us to reconfigure our narrative identity by freeing us from unwanted narrative constructions brought about by information processing. In such information processing, technology and human agents both play a role: technological intentionality and human intentionality are both a part of a hybrid affair that comes into being with the use of a technology. However, while art. 17 GDPR certainly is of importance vis-à-vis the presence of narratives that for the majority result from human intentionality, I find it important to especially highlight the potential functionality of art. 17 GDPR to deal with the impact of technological mediation on the narrative identity, and especially the impact resulting from a too strong expression of technological intentionality. While technological intentionality is always a part of a hybrid affair of technology and its user, giving extra attention to the technological intentionality expressed in the concrete information processing can help to resolve problems in a fine-tuned and balanced manner.

In this guise, art. 17 GDPR can function as a counter option to the forms of processing that dominate the shape of the materialised narrative identity, and can thereby be employed as a *counter technique*. With 'counter technique', I refer to a manner of counteracting the impact of specifically technology, and in particular in the direction of its intentionality. The right should aim to counter the effects of the dominating technology by means of erasure. Especially paragraph (2) of art. 17 GDPR, which aims to account for the implications of the multiplication and transmission affordances of online digital information, makes sense when we approach the right in the perspective of aiming to function as a counter technique

in the internet era.[8]

The help offered by art. 17 GDPR to reconfigure problematic materialised narrative identities needs to be shaped on the basis of a balance of interests. For this balance of interests it is helpful to also consider art. 17 GDPR in the context of Agre's view on the right to privacy and identity. Agre defines the right to privacy as "the freedom from unreasonable constraints on the construction of one's own identity" (Agre, 1997, p. 7). In this definition the negative freedom, a freedom from constraint, is combined with a positive freedom, a freedom to build our identities (Hildebrandt, 2006). I agree with Hildebrandt that it is particularly important to understand privacy rights at the axis of this negative and positive freedom. The freedom *from* something, is often a requirement to actually realise the freedom *to* something. For instance, when my online narrative is fully plotted in a manner that defines me as an X, I may have a freedom to define myself as a Y, but this freedom is only fictitious if I am unable to free myself from this narrative impressed upon me by the outside world. This example also shows that positive and negative freedom together can give rise to an interplay that allows a nuanced give and take: I do not need to be free from all objects containing a reference to X in order for myself to present myself as a Y, I just need to be free from being *overruled* by X. This connects to Agre's focus on protection from 'unreasonable constraint'. With this, Agre poses the right to privacy not as an absolute right, but instead places it in a broader context that leaves room for a balance of interests (Hildebrandt, 2006). In accordance with this, the wide scope of erasure offered by art. 17 GDPR fits this perspective as the right offers a sliding scale of erasure applications to free oneself from an unreasonable impact of technological or human others on the construction of our narrative identity. Whether the right should be granted, and in what form, depends on the impact that the particular account of information processing has on our identity building, as well as on the interests of others.

Given the balance of interests and taking the problem analyses of the previous chapters into account, I thus argue that, deployed as a counter technique, art. 17 GDPR should be aimed at reducing the presence of a particular reference in accordance with its accuracy and proportionality viewed in relation to the manner in which the processing of the personal information affects the narrative identity on the level of the narrator, the plot and/or the composition of audiences. Ideally, the right to erasure should have the upper hand in cases where the technological intentionality shapes the narratives beyond human storytelling, intentions and expectations. This approach places the focus on the value of *human* autonomy and dignity, by taking human intentions as an important guiding principle for assessing whether a certain technologically mediated portrayal of an individual should be addressed or not (of course, a strong expression of human intentionality is not a free pass: content that is processed under a strong human intentionality could also be a potential target for erasure). It also ties in with two important points of the GDPR: its rationale and its focus on the purposes of the controller. By placing

[8]It is important to underline here that making sense conceptually is something different from being effective.

human intentions centre stage, understanding art. 17 GDPR as a right that can be deployed to address an overly strong impact of technological intentionality is in line with the GDPR's rationale that the "processing of personal information should be designed to serve mankind" (recital 4 GDPR). Additionally, by taking human interests as a guiding principle, this approach connects to the GDPR's overall focus on the goals of the controller: as the legality of the information processing depends on the purposes of the controller, her intentions with regard to the processing of the information play a pivotal role in the balance of interests. In this balance, the respective influence of the controller and the technology in the hybrid intentionality should be taken into account. A claim to art. 17 GDPR would especially carry weight in cases where the mediating technology increasingly impacts the narrative.

9.2.5 Looking at the cases

With this initial framework outline, I have shown that the construction of the externalised narrative identity is a hybrid affair. The technological mediation of the information affects the trajectory of the narrative by pressing its own intentionality on the narrative act. This frame provided a foundation for me to sharpen the conceptualisation of art. 17 GDPR in a manner that will allow me to evaluate the right in the context of the three cases of technological mediation, as well as the phenomenon of virality, that I investigated in chapters 4 to 7. I will do this in the upcoming four sections. Based on this analysis, I will assess the problem-solving potential of art. 17 GDPR and identify its strengths and weaknesses. For clarity purposes, I will split up the application of art 17. GDPR in three main actions: invoking the right, balancing the interests, and executing the erasure. Depending on the case, some of these actions will be intertwined in the actual application of the right (for example, invoking art. 17 GDPR on ground (c) already contains a balance of interests because the controller has to assess whether she has overriding legitimate grounds to retain the content).

With regard to the analysis in this chapter, a remark is in order. The cases investigated in chapters 4 to 7 are highly dynamic in the sense that their mechanisms are constantly adjusted and updated (maybe even in response to issues like those discussed in this study). The analysis of the application of art. 17 GDPR to these cases therefore sees to a particular moment in time and functions as a rough exemplary blueprint for how to ideally approach the application of art. 17 GDPR in certain types of contexts and in relation to the main informational persona affecting mechanisms identified in these cases. In case of changes to the software architecture and its mechanisms, the framework can provide handholds to adjust the analysis for a particular case.

9.3 Web pages

When mediating our representation, the Web's character leaves an imprint on the narratives that it presents. We can already see this on the level of the 'who' that can become an online narrator. By allowing virtually everyone to publish for a potentially global public, at the expense of little effort, the Web gives a voice to a vast array of different narrators. Many of these use the Web for social interactions and as a means of self-expression. In the early years, users often expressed themselves on personal home pages (de Mul, 2002, p. 211). The flexible affordances of digital information allows users to tinker endlessly on their web page to shape their identity. The 'old school' personal home pages therefore often consist of materialisations of the referent's self-identity in progress (de Mul, 2002) — that is, unless someone forgets the page or loses access to it, and the portrayed plot remains fixated over time. Currently, online self-expressions have mainly moved to social media — which give quite a different twist to the materialised presentation of the self in progress. I will discuss this in the next section. In this section, I will keep the focus on web pages in general.

When personal references are encoded online, they become a hybrid affair of a human and technological narrator. Once online, they become part of the informational persona of a referent and are absorbed in the narrative identity that is constructed by the presence of the diverse objects that constitute the informational persona. However, the total of this online narrative identity is rarely narrated by one hybrid narrator set: the references that shape the informational persona can be encoded in signifying objects by the subject herself, as well as other users and third parties. Online, people can easily become (co)biographers of each other's narrative identity (Buitelaar, 2014, p. 274). This can happen with or without consent of the referent. As a publishing area open to a wide and mixed array of narrators who can publish at any time, the Web easily gives rise to an extensive informational persona consisting of many and diverse references. Users publish information about each other, reproduce and remix each other's content, etc. Additionally, due to the affordances of digital information, personal signifying objects are easily hijacked, changed and/or decontextualised by users and by technologies (such as search engines) alike. Even objects that are processed with consent of the referent can lead to problems, because the referent may overlook or misjudge the implications of the particular instance of online processing. As the objects collide with each other and interweave online, they affect the narrative identities that they present. With this mix of different actors and objects, the construction of online narrative identity is thus rarely the sole activity of the individual herself. Instead, it takes shape in a mix of actions performed by the referent, other users, and the mediating technology. This brings me to the question of how these narrators together, albeit not necessarily with an equal impact, construct an online narrative identity. For clarity's sake, let us trace this from the beginning so that it can serve as foundation for the following sections.

Online, referents are constituted as a subject of particular plotlines that tell a certain story about the permanence and change of their identity. This plot is

shaped by the manner in which online signifying objects constitute the presence of certain personal references. A first narrator publishes a particular signifying object, which possibly already contains quite an extensive narrative content, on a website. This content in the signifying object is the most basic building block of the online narrative identity. However, this object is rarely 'alone': websites often display multiple objects in a mixture of images, text, videos and possibly ads. As such, websites come across as a kind of informational collage (de Mul, 2002, p. 213). The other content on a web page, as well as the page's header and URL, shape the context of the signifying object and thereby affect its meaning.

Moreover, as the object is absorbed in the networked weave of the web, it is embedded between other pages and becomes open for hyperlinking. The narrative presented by the object is embedded into the overarching plotlines afforded by the Web: the story told by the associations made by hyperlinks and adjacent content. The Web imprints its character on the narrative by allowing it to be accessed by an audience from anywhere at any time, thereby affecting the context in which the story is read, while the story takes shape along an associative path across the Web's hyperlinks. The narrative thereby inherits the affordances of its technological narrator and can be shaped over and over again in myriad ways along the trail of hyperlinks (de Mul, 2002, p. 213). The result is that the online narrative is dynamic, consistently on the move and subject to new cycles of retrieval, disclosure, dissemination, combination and collisions with other online narratives. The main configuration of the overarching online narrative of a particular data subject as presented to users revolves around their interests, as they follow these through trails of hyperlinks. This allows each user to get a different view of the data subject's narrative identity. However, currently, this navigation is often affected by search engines. I will discuss the implications of this for the online narrative identity in section 9.5.

While the online narrative identity is highly flexible and spread over different sources, it at the same time can give rise to a certain persistence of particular plotlines. This is where the online narrative identity is imbued with a certain sameness of character, a part of the informational persona that is unchanging over time. The objects that have a high and longterm presence can become the lasting dispositions by which the referent is recognised by others, as well as by herself. The question is whether the plot oscillates between sameness and change in tune with the referents' embodied sense of personal identity, or whether there is a fundamental dissonance between the narrative that the Web presents and the narrative that individuals (hope to) maintain in their lives. For certain aspects of the informational persona, like a name, this generally static presence of the information is often not a problem. However, when there is a significant dissonance between the referent's externalised narrative identity and her sense of self, this can clash with her self-perception and lead others to respond to her in ways that are incompatible with her.

The audience is another element of online narrative identities that can raise problems. As part of a hybrid narrator set, the Web heavily impacts the composition of the narrative's audiences by offering continuous global access to the

narrative objects. Once problematic content becomes part of the online narrative, it is therefore difficult for individuals to escape the audiences of these plotlines. Especially given the Web's dominant role in the fabric of contemporary society and our daily praxis, online narratives can have a strong presence for a potentially large audience.[9] The human narrator is therefore confronted with a challenging task when narrating to online audiences: she faces a culturally undetermined and faceless audience, while having difficulty to control the context of the story that she narrates. She is left to the realm of digital code to express herself, and may never know if she has an audience and whether it is her intended audience. Audience segregation faults (see section 2.3.2) are therefore easily made, especially given the fact that plain web pages tend to be publicly accessible by default.

Altogether, the Web potentially gives rise to an externalised narrative identity with a scale, scope, flexibility and access speed that can easily result in a longterm and global presence. Online narrative identities can become a personal billboard that is synchronously present next to offline and other online interactions. The effects of the online architecture ripple through in every detail of these narratives; with the many players and informational building blocks, the Web gives rise to interactive narratives to which anyone can contribute, which can appear in unintended and unforeseen manners, and which can reach equally unintended or unforeseen audiences. Meanwhile, the problems of the virtual narrative can have real consequences. For example, the presentation of outdated information on a website, like photos of a happy marriage that in fact has ended in a divorce, may easily leave a wrong impression of the data subject with potential new dates who are unaware of her divorce. As others respond to the referent based on this information, certain options are opened or closed to her, while the reactions of others also reflexively affect the referent's self-perception. Additionally, being confronted with the content herself can affect the referent's self-perception by, for example, triggering dormant memories or providing her with an unexpected view of herself.

9.3.1 Applying art. 17 GDPR to web pages

In this section, I will discuss the application of art. 17 GDPR to basic websites. This analysis will serve as foundation for the following ones. Because art. 17 GDPR targets signifying objects (including complete datasets) or their descendant objects, I find it worthwhile to first provide some structuring demarcations and distinguish between six general situations in which a signifying object can be added to the Web[10]:

- (I) The subject herself encoded the content online, and is also the controller

[9] Yet, the actual presence of online information will always depend on the engagement and choices of the individual user and online applications like search engines.

[10] Not every real online situation will perfectly match the situations listed here. In the case that an online situation is a mix of two or more situations here, or only partially resembles a situation, one can take into account the relevant aspects of the general situations and use this as a guideline for the application of art. 17 GDPR.

of the content[11];

- (II) The subject herself encoded the content online. However, the web page on which the content is uploaded is under the control of someone else (e.g., the subject comments on a news website that does not allow her to edit the content once it is posted);
- (III) Someone else encoded a descendant object of content that was initially uploaded by the subject. This descendant object is under the control of this other;
- (IV) Someone else encoded a descendant object of content that was initially uploaded by the subject, but the descendant object is under the control of yet someone else;
- (V) Someone else published a new object about the subject and is the controller;
- (VI) Someone else published a new object about the subject, but the object is under the control of yet someone else;
- (VII) A mediating technology publishes a new object about the subject (e.g., Facebook publishing "Captain Picard liked this picture"). I will assume the controller of the website is in these cases also the controller of these automatic publications;
- (VIII) A mediating technology publishes a descendant object of the objects published in (I-VII).

For completeness' sake, I have added situation (I), but as I will quickly explain, this is a non-issue for art. 17 GDPR. Furthermore, problem cases can cover multiple situations, and some of the situations are a necessary condition for other situations. For example, if the problem is that the content of situation (I) is copied and stored in an archive or the like, then the case concerns situation (III), but it could not exist without the occurrence of situation (I). Additionally, it is important to remark that situations (III) to (VIII) could in theory transpire in two versions: (a) with the data subject's consent, and (b) without the data subject's consent.

Situations (VII) and (VIII) will be discussed in detail in sections 9.4 and 9.5. In this section, I will focus on (I) to (VI) and walk through the three main actions involved with applying art. 17 GDPR in these cases: invoking the right, balancing the interests, and applying a form of erasure.

[11]The subject may have a joint controllership. Because joint controllership means that the controllers jointly determine the means and purposes of the processing, the controllers together initially will need to work out what to do when one of the controllers wishes to erase the content and the other not. If they disagree and the data subject does not succeed in having her information erased, we can question the extent to which she still can be considered to be a controller in this case. She then might be able to request a right to erasure, and I expect the case will roughly follow situation II.

9.3.1.1 Invoking art. 17 GDPR

As discussed in the previous chapter, art. 17 GDPR is subject-driven. In order to adjust or correct the online personal narrative with the use of art. 17 GDPR, an individual needs to take action by addressing her request for erasure to the controller of the signifying object. In situation (I), the subject is the encoder *and* controller of the medium and object. Practically, invoking art. 17 GDPR against herself is futile, as she can erase the content herself at will. Theoretically, it is questionable that a data subject can even be marked as a data controller in the legal sense with regard to her own personal information, because imposing all the obligations of a controller on an individual who only processes information about herself seems disproportionally burdensome. While the subject has no use for art. 17 GDPR in situation (I), the content can still raise problems. For example, an individual may believe that she is publishing information about her holiday for a limited group of friends, but by mistake she also shares the information with unintended audience members, like burglars. However, because the website is under control of the data subject, she can erase the content herself and has no need to invoke art. 17 GDPR. This shows that art. 17 GDPR does not resolve issues that transpire under the control of the individual herself. I will therefore not consider (I) further for the other cases. I take situations (II) to (VIII), where the subject of the information is not also its controller, to be the target of art. 17 GDPR. In these cases, art. 17 GDPR needs to be invoked against a certain other.

In order to invoke art. 17 GDPR against a certain other, the data subject will first need to identify who this controller is, and how to reach her. If the contact information of the web page controller is listed on the page, this is not a problem. However, when there is no controller information listed, the individual is dependent on secondary sources in order to identify the controller. Examples of such secondary sources are organisations that keep track of the owners of domain pages. The accessibility of this information differs per domain name and country. While there used to be publicly accessible databases for this, like the WHOIS protocol that can be used to retrieve information about registered domain names (Daigle, 2004), this information has become less accessible over time. Data protection laws like the GDPR play a role in this. If the data subject does not succeed in identifying the controller though such secondary sources, she can try to get the information from the Internet Service Provider (ISP) hosting the web page. Given that these providers are also subject to data protections laws, it is unlikely that they will hand over information about their customers easily. An example of this is the Dutch Lycos/Pessers case. In this case, a client of the ISP Lycos anonymously accused a data subject of fraud on his website. The data subject turned to Lycos to request the personal information of the website controller. Lycos refused, and was brought to court by the data subject. In the end, the Dutch Supreme Court ruled that the ISP was required to release personal information of its clients *only if* the content is unmistakably unlawful.[12] The standard for revealing personal information by ISP's is thus high. Because art.

[12]Dutch Supreme Court, *Lycos/Pessers*, 25-11-2005, ECLI:NL:HR:2005:AU4019.

17 GDPR should ideally apply in a far broader range of cases, like cases where content is merely outdated or disproportional, the Lycos/Pessers standard could be a significant stumbling block for its application. However, it is important to remark that this is not a limitation of art. 17 GDPR itself and this obstacle should be addressed elsewhere in law.

Furthermore, in order for the data subject to be able to successfully invoke her right to erasure, the controller of the targeted content needs to fall within the territorial and material scope of the GDPR. Given the character of the content that is the focus of this study, I will restrict the discussion of the material scope to the household exemption (art. 2(2)(c) GDPR). However, let me first briefly touch upon the territorial scope. While it lies outside the scope of this study to map the exact impact of the GDPR's territorial scope (art. 3 GDPR) for the application of art. 17 GDPR, I expect that with regard to signifying objects on regular web pages, some of the cases are likely to fall outside this scope. The reason for this is that web page controllers can originate from anywhere and do not necessarily target an EU audience while processing content of EU data subjects (see section 8.2.4). In contrast, the household exemption will have little impact on art. 17 GDPR's reach with regard to many basic websites. To begin with, publicly accessible websites do not fall under the household exemption. In the Lindqvist ruling, the CJEU ruled that publishing personal information about volunteers of a church community on a publicly accessible website does not fall under household use.[13] The reason for this is that on such a website, the content can be viewed by an indefinite number of people.[14] The consequence is that publicly accessible websites, even if controlled by private persons for nonprofessional use, fall within the working range of art. 17 GDPR. Furthermore, the information processing on a significant part of the more closed off websites are also unlikely to enjoy protection under the household exemption because their audiences may easily exceed the household use scope. I will discuss the scope of the household exemption in relation to restricted websites further in section 9.4.

Lastly, the data subject will need to base her request on certain grounds. With regard to web pages, grounds (a) (the information is no longer necessary), (b) (the subject withdraws consent), (c) (the subject objects to the processing), (d) (the information is unlawfully processed), and (e) (the information needs to be erased for legal compliance) can be appropriate grounds, depending on the specifics of the case. However, given the global character of the Web, it is important to mention that ground (e) may give rise to some issues as laws on retention and erasure can differ per EU country, and these laws may clash. Moreover, while for ground (d) the GDPR has harmonised the information processing rules within the EU, this ground may be less effective in cases where the website controller is located outside of the EU, but "in a place where Member State law applies by virtue of public international law" (art. 3(3) GDPR). In this case, the web page controller might be subject to different processing laws that allow her to lawfully process the

[13] CJEU, 06-11-2003, C-101/01, ECLI:EU:C:2003:596 (*Criminal proceedings against Bodil Lindqvist*).

[14] Ibid., §47.

information — despite it being unlawful in the EU.

If the primary conditions for the erasure request have been fulfilled and the controller falls within the GDPR's territorial scope, the data subject's can make a valid request for the erasure of specific content on a web page.

9.3.1.2 Balancing the interests

Another pivotal element of the application of art. 17 GDPR is the balancing of interests. The requirement of such a balance is embedded in several parts of the right. The invoking of the right on ground (c) contains a balance of interests because the controller has to assess whether she has overriding legitimate grounds to retain the content. Additionally, ground (b) may entail something of a loophole for the controller to balance her interests, if she can argue that there are other grounds of processing besides consent. However, the most evident balance of interests that is embedded in art. 17 GDPR is paragraph three, the exceptions to the application of the right. Because discussing all the varieties of the balance of interests in detail is a study on its own and will differ highly per case, I will restrict myself to sketching a general balance of interests and discuss what I take to be the main points with regard to online content. In this subsection, I will therefore sketch a general picture of the various factors that (should) play a role in the balance of interests. I will raise some questions and offer some suggestions on how to approach the balance of interests in the case of an art. 17 GDPR request. Because web pages are the most basic version of online information sources I discussed, the balance of interests discussed here will function as the base for the other case analyses in the next sections.

With regard to content on regular web pages, the main reasons to not apply erasure are when the processing of the information is necessary (a) for exercising the freedom of expression and information, (b) for compliance with a legal obligation, and (d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. Exceptions (c) and (e) are relatively unlikely to be used to legitimate the retention of information on web pages in a publicly accessible manner (see section 8.2.9). I will therefore leave these outside the further analysis with the remark that if these cases occur, the general points of the balance of interests that I will raise here can be used as a backdrop to assess the proportionality of the erasure request in relation to the particular interest that is covered by the provision.

For practical reasons, I will start the discussion of the balance of interests with the exception that will likely leave the least room for actual balancing: exception (b), when the processing of the personal information is necessary for compliance with a legal obligation. An example of this is the original publication of the newspaper articles by *La Vanguardia* in the *Google Spain* case. *La Vanguardia* was legally obliged to publish this content in the newspaper. However, this raises the question of whether the obligation to publish content in a newspaper also entails the obligation to publish it online and retain it there indefinitely. This will depend on the (interpretation of the) law in question and I cannot presume

to answer this question in general or rather (here), for Spanish law. However, I argue that lawmakers as well as courts thoroughly need to consider what the scope of the intended audience is, and whether an indefinite online publication is necessary to fulfil these legal obligations. Additionally, this case raises the question of whether the legal obligation to process and publish information also stretches to the archiving of information, and even more, to the online publication of the archive. I doubt these obligations in all cases stretch this far. However, — and this is where I get to exception (d) — archives themselves are also protected.

Exception (d) is a relevant exception with regard to websites, as many information sources like newspaper agencies, musea, and governmental institutions offer their archives online for others to use for purposes in the public interest, scientific or historical research purposes or statistical purposes. First of all, the balance of interests with regard to the erasure of information from online sources that process information for archiving purposes, is complicated. Two European Convention on Human Rights (ECHR) cases that are relevant to touch upon here, are the case of *M.M. v. the United Kingdom*[15] and the case of *M.L. and W.W. v. Germany*[16]. In the case of *M.M. v. the United Kingdom*, an individual sought to have information erased from the public record about a police caution that she received in 2000. In this case, the court decided that the indefinite retention and public availability of information concerning a police caution that was received by the data subject, was in breach with art. 8 of the ECHR (the right to respect for private life). The court states that "although data contained in the criminal record are, in one sense, public information, their systematic storing in central records means that they are available for disclosure long after the event when everyone other than the person concerned is likely to have forgotten about it, (...). Thus as the conviction or caution itself recedes into the past, it becomes a part of the person's private life which must be respected".[17]

In the case of *M.L. and W.W. v. Germany*, the court reaches another decision. The cases revolves around two individuals who are convicted of murder. After they sat out their punishment, they wanted to have the information relating them to the murder removed from newspaper archives. M.L. and W.W. did not request the full erasure of the information, but only their anonymisation. While the court acknowledged that this was a less restrictive measure than full erasure, it decided that this choice was a matter of journalistic freedom and that there was no violation of art. 8 ECHR.[18] Several factors played a role in this. First of all, the case in itself has a very different character from that of *M.M. v. the United Kingdom* (i.e., the difference between a caution and murder). However, what also played a relevant role was M.L. and W.W. requested the reopening of the case, and also contacted the press on this matter in 2004. The court argued that "as a result of the applicants' conduct vis-à-vis the press, less weight was to be attached to their

[15]ECtHR, 13-11-2012, application no. 24029/07 (*M.M. v. the United Kingdom*).

[16]ECtHR, 28-06-2018, application no. 60798/10 and 65599/10 (*M.L. and W.W. v. Germany*), press release 237 (2018), 28-06-2018.

[17]ECtHR, 13-11-2012, application no. 24029/07 (*M.M. v. the United Kingdom*), §188.

[18]ECtHR, 28-06-2018, application no. 60798/10 and 65599/10 (*M.L. and W.W. v. Germany*).

interest in no longer being confronted with their convictions through the medium of archived material on the internet. Their legitimate hope of obtaining anonymity in the reports, or even a right to be forgotten online, had thus been very limited"[19].

The balance of interests with regard to public archived information sources, is a difficult balance with regard to the application of art. 17 GDPR. I argue that two points are relevant to consider for this balance. The first point is that we should consider who the intended audience of the archive is and whether a global online accessibility really reflects the intended audience. Secondly, I argue that in these cases we should especially consider the role of the technology and its impact when we balance the interests. What is important to underline in this context, is that the exceptions "apply to the extent that *processing is necessary*" for the named purposes [my emphasis] (art. 17(3) GDPR). If we then look back at art. 4(2) GDPR, we can see that processing entails "any operation or set of operations (...), such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction". The exception to processing therefore does not necessarily mean an exception to *disclosure by transmission, dissemination or otherwise making available on the Web*. Offline collection, storage and dissemination of information are equally plausible manners of processing. The question therefore is, if the term 'processing' in art. 17(3) should be understood as all-encompassing, meaning that once something falls under the exceptions, all manners of processing are allowed, or if it needs to be understood in proportion to the purposes of the processing. Given the rationale of the GDPR, I argue that the latter is the case. Additionally, this seems to be also the view of the European Archives Group.[20] One of the points that they raise in relation to the implementation of the GDPR in the archive sector is that "[s]toring personal data is not the same as providing access"[21]. In general, many archiving institutions keep personal data, especially sensitive personal data, closed for 40 to 100 years after their creation.[22] Given the fact that the GDPR does not apply for deceased persons, such closing periods of collected documents will reduce the chance of archived content becoming a target of an art. 17 GDPR request. In those cases that apply to a living data subject who experiences the content as problematic, art. 17 GDPR can be of help as a tool that supports proportional processing and dissemination of information in line with the purposes of the archive. With regard to institutions that make their archives accessible online, the European Archives Group recommends the following:

whenever [archivists] post archival documents or finding aids that contain personal data of living individuals online, they have to consider — according to

[19]Ibid., press release 237 (2018), 28-06-2018, p. 3.

[20]European Archives Group, *Guidance on Data Protection for Archive Services* - EAG guidelines on the implementation of the General Data Protection Regulation in the archive sector, `https://ec.europa.eu/info/files/guidance-data-protection-archive-services_en`, last accessed 03-09-2019.

[21]Ibid., p. 6

[22]Ibid., p. 6.

the nature of the personal data — whether it would be appropriate to post them in a restricted-access area of their websites which is out of the reach of search engines. On a case-by-case basis, archivists will assess how to best balance their legal obligation to 'describe, communicate, promote, disseminate and provide access to records of enduring value for general public interest' (recital 158) with the principle of data minimisation (art. 5), which requires them to limit the processing of personal data to what is necessary.[23]

This recommendation responds to the impact of an archive being absorbed by the Web and thereby becoming open to online applications like search engines. As such, the European Archives Group thus argues in favour of the website controller (in this case the archive controller) taking responsibility to realise proportional online processing also with an eye on what other online applications do with the content. In case of a disagreement between a data subject and an archive controller on the availability of particular content, the wide scope of art. 17 GDPR's erasure can be of help to address (a part of) the issue in the case of online archives, without necessarily requiring the complete erasure of the content. I will discuss some possible uses of partial erasure in the next subsection.

We now arrive at the next, and most complex, exception to art. 17 GDPR, namely exception (a): the processing is necessary for exercising the right to freedom of expression and information. I already touched upon the general value, importance and limitations of the right to freedom of expression and information in section 8.2.9.1. What is important to consider here, is that not every speech act is given equal value (e.g., political speech enjoys a stronger protection than commercial speech, see section 8.2.9.1). Also, the exact extent of the protection can differ per country. In this section, I will not go into a detailed weighing between art. 17 GDPR and the freedom of expression because this depends heavily on the circumstances of each case. Instead, I will focus on the crossroad between freedom of expression, the right to erasure, and the impact of the technological mediation. With an eye on this crossroad, I will suggest a manner to approach the balance of interests in cases that appeal to the freedom of expression and information.

In this balance of interests with regard to public websites, we can identify three stakeholders: the controller, the data subject, and the general public. However, if we look back at the situations in section 9.3.1, we should note that there is a pivotal difference between the role of the stakeholders in situation (II) on the one hand, and (III) to (VIII) on the other hand. In situation (II) the data subject herself is the initial publishing narrator of the content, while in the case of (III) to (VIII), the publishing narrator is another party. While this is certainly an important difference with regard to the balance of interests, I argue that this difference is of a lesser importance, the more the mediating technology leaves its imprint on the narrative. I will clarify this in the following paragraphs. I will first discus the balance between the data subject and the expresser in the light of technological

[23]European Archives Group, *Guidance on Data Protection for Archive Services* - EAG guidelines on the implementation of the General Data Protection Regulation in the archive sector, p. 18.

mediation. Next, I will add the general public to the mix. This will be followed by a discussion of the roles that various kinds of content can play. I shall conclude this section with an overview of the diverse factors that play a role in the balance of interests and their respective weight in favour of or against erasure.

The intentionality of the narrator If we want to do justice to the problems raised by technological mediation, I suggest that an important focal point of the balance of interests should lie in the degree to which there is a *direct connection between the human intentionality of the controller and the presence of the reference.* First of all, this ties in with the purpose specification and purpose limitation principles of the GDPR because the core of the processing revolves around the purposes determined by the controller. As WP 29 already accurately pointed out: "[d]ictionaries define 'purpose' as 'an anticipated outcome that is intended or that guides your planned actions'"[24]. The purpose limitation principle is thus rooted in the intentions of the controller.

Secondly, and maybe even more relevant for the balance of interests with regard to art. 17 GDPR, valuing this connection does justice to the values attributed to the freedom of expression (I will discuss the value of the freedom of information in the next subsection). In section 8.2.9.1, I briefly discussed the value(s) attributed to the freedom of expression. In this, a certain weight is given to the freedom of the individual to express her ideas and feelings. This is a view that we find clearly expressed in the Dutch constitution where the freedom of expression literally entails the freedom to make *thoughts or feelings* public (art. 7(1) Gw). The expression of thoughts and feelings requires a certain intention of the expresser, and more particularly, a directionality towards specific thoughts and feelings. The content has a certain meaning to the expresser, and she has a reason to want to publicly express it. I therefore argue that the human intentionality plays an important role in the freedom of expression. Yet, I argue that in the case of technologically mediated publication, the human intention can be watered down, even to the point where the 'expression' is the result of the technological intentionality. An example of this is the automatic publication by a social media platform that refers to the action of a user, like "Paulan likes Zerum". Because human beings are the pillar of freedom of expression, I suggest approaching the degree of the protection under the wings of freedom of expression in proportion to the role of the human thoughts and feelings in the constitution of the presence of a particular reference. Let me explain this approach by discussing it in relation to the balance of interests in situations (II) to (VIII). Although situations (VII) and (VIII) will rarely occur in the case of basic websites, I already briefly touch upon this here for clarity's sake. In the following sections on social media and search engines, I will discuss these situations in more detail.

Starting with situation (II), which sees to a signifying object created by the data subject herself (albeit in a hybrid act with technology), but once published, the object is under control of someone else. In this case, we can generally understand

[24] WP 29, Opinion 1/2010 on the concepts of "controller" and "processor", p. 13.

the publication of the initial object as an intentional expression of the subject's own narrative identity: the subject herself establishes a narrative with her personal information. However, it is important to note that with the use of increasingly easy publication mechanisms and the like, this human intentionality may in some cases be on the weak side: some of her publications may be the result of unthinking, hasty push-button-publishing actions. If the content is not the result of a spur-of-the-moment action that the subject later regrets and the subject intentionally creates the object in a narrative setting by her own choice, the problems are likely to be caused by (1) narrating to an unforeseen or unwanted audience, or (2) by an unforeseen persistent presence of the object due to which it may establish as an undesired expression of sameness in the individual's online narrative identity that persists over time. In both cases, the data subject cannot erase the content, because the website is under the control of someone else. This controller thus has the control over the presence of the signifying object over time, and therefore has a certain control over whether an expression of the referent may in time become persistent and a thereby a longterm portrayal of the referent in her materialised narrative identity. However, given that the object's presence over time is always the result of a hybrid intentionality, the mediating technology also plays a role in this. I therefore argue that it is important to consider the relation between the intentionality of the controller and the mediating technology.

In some cases, the controller may have an interest in retaining the signifying object on the website, and thus does so intentionally. As such, the intention of the controller plays a fundamental role in the presence of the reference over time and should therefore have a significant weight in the balance of interests. However, in other cases, the controller may pay little attention to that specific part of the website or may even have forgotten it and retains the content merely because it is stored by default. In this instance, the technological intentionality has the upper hand in establishing the presence over time because the storage by default is the key factor in retention of the information, while human intentionality plays a secondary and minor role. I therefore argue that in such cases the interests of the controller should be given less weight compared to situations where she clearly expresses a certain intention by for instance updating the content or hyperlinking to it. Hence, I take the source and the intensity of the intentions that shape the narrative as important factors in the balance of interests. The focus of art. 17 GDPR on 'exercising' the freedom of expression and information ties in to role that I attribute here to intentionality because it underlines the weight of freedom of expression as a human action, and not as a merely technologically mediated passive existence.

Situation (V) and (VI) are probably the most 'classic' freedom of expression cases and result from a human intention to express oneself. As such, these situations should receive commonly the greatest protection under the freedom of expression of cases (II) to (VIII). However, as these narrative acts take place on the Web, the impact of the mediating technology on the composition of the audience and the development of the role of the object in the plot of the subject's narrative identity over time, should also be taken into account.

In the case of situations (III) and (VIII), the data subject is the initial narrator, but she is not the sole narrator. As the signifying object is assimilated by others (human or technological agents) into the narratives that they actively publish, these agents take on the role of narrator$_2$. By republishing the signifying object in another context, these agents press their own intentionality upon the narrative plot. This may result in unforeseen or unwanted effects on the narrative identity of the data subject. Here, again, I argue that we should differentiate between the extent of the human and the technological intentionality for the level of protection. The consequence of this is that in the cases where another human agent encodes (descendant) content, this content should in general receive a greater protection under the freedom of expression than in cases where the mediating technology is the driving force in the publication. However, it is important to note that this will always entail a sliding scale. For instance, in the case of (III) and (IV), there is a sliding scale between human and technological intentionality due to the current highly interconnected character of the Web. Online, many applications are offered that allow users to republish content with a single click of a button, requiring little time to reflect or think. In such single-click republications, the expression of human intentionality can be weak, while the expression of the technological intentionality is relatively strong — including nudges towards users to republish content (I will get back to this in section 9.4). The role of human thought will differ highly per case, and will in all likelihood make the balance of interests a difficult evaluation. Despite this difficulty, I argue that for a just and beneficial application of the right to erasure, the role of human intentionality is important to take into account.

Lastly, it is important to remark that (IV), (VI), and (VIII) entail a somewhat different balance of if interests than (III), (V), (VII). In the case of the latter, the narrator is also the controller, while in the case of the former, a third party processes a narrative uploaded by someone else. I expect that invoking art. 17 GDPR vis-à-vis the uploader-controller might be simpler than vis-à-vis a controller who did not publish the object. This is because in the case of a third party controller that processes content uploaded by others, the interests of the controller as well as of the uploading party need to be taken into account and the controller has to balance the freedom of expression of the publisher as a separate (and weighty) interest.

Overall, when looking at basic websites from the perspective discussed here, we can conclude that content on basic websites is generally subject to a relatively strong intentionality of the expresser. I therefore argue that content on basic websites should receive a substantial protection by the right to freedom of expression (it is important to note that other factors, like the character of the content, may still overrule this protection, I will discuss this later). However, this protection should be watered down with the loss of the relative weight of their intentionality in retaining and/or distributing the content as the mediating technology expresses a strong intentionality. In the cases where the intentionality of the mediating technology has the upper hand, the interests of the data subject should receive significant weight compared to that of the expresser and controller. These are the

cases in which the content is retained while the expresser or a mediating controller does not have any significant interest in this retention. An ongoing retention of the content in such situations should be based on something other than the expresser's and controller's (lack of) interests, namely the interests of people in gathering information. I will discuss this in the next subsection.

Freedom to gather information In the previous subsection, I have discussed only one half of the value of the freedom of expression and information, namely the value for the expresser. I shall now look at the other half: the freedom of individuals to gather information. As a public communication network, the Web serves people by providing access to information. However, the affordances of the Web highly affect the scale and scope of the public that can access information compared to offline publics. I argue that it is important to think about the scope of the public that has a right to gather this information. Two things that are worthwhile to consider with regard to the scope of the public that has an interest in accessing particular information, are (1) the relation between the public and the content, and (2) the interest of individuals in this information over time.

Starting with the first. Content is generally imbued with a particular lingual, cultural and/or national character. Yet, the networked character of the Web does not necessarily follow this composition of the public and may easily disclose the information to publics with different identity characteristics. The disclosure of information to unforeseen publics, even when it is intended to be public, can give rise to cultural misunderstandings, prejudices or even condemnations that were not part and parcel of the intentions behind the publication. For example, a woman may have given an interview to a Dutch LGBTQ (Lesbian Gay Bisexual Transgender Queer) magazine about her experiences as a lesbian in order to support and inform LGBTQ youth. However, imagine that at a certain stage, her job requires her to work with local authorities of a country that is strongly homophobic and upon discovery of the online interview, the authorities of this country demanded that she is replaced as a contact person.[25] While these authorities are far from the intended audience of the online magazine, due to the affordances of the Web — especially with the use of a search engine — this information can become relatively easily accessible to such unintended audiences, with ensuing consequences. I therefore think that it is worthwhile to consider the scope of the group of individuals that has an interest in gathering particular information in a contextual manner (cf. Nissenbaum, 2004). If we take this into account as a factor in the application of art. 17 GDPR, I can imagine that in some cases it may be worthwhile to consider restricting the processing to particular publics in response to an art. 17 GDPR request. I will get back to this in the next section where I discuss the various forms of erasure. Restricting information access to a public of which we can reasonably expect that they understand the content in a correctly contextualised manner, could allow us to reduce the impact

[25]This is a fictive example, but inspired by an actual incident experienced by a data subject that I spoke to.

of the Web's mediation, without disproportionally harming the right of individuals to gather information — but only if done with care.[26]

Additionally, the factor 'time' plays a role in the evaluation of the interest that individuals may have in certain information. In his paper *The right to be forgotten: balancing interests in the flux of time*, Sartor argues that in the evaluation of the balance of interests, the factor 'time' should play an important role as it affects, and can even flip, this balance (Sartor, 2015). As people change and evolve over time, the relevance of particular information as a reflection of who they are now, can reduce to the point where the interests of the individual to have the content erased, weigh more heavily than the interest of the public in the retention of the content. According to Sartor, an art. 17 GDPR request should therefore not be a question of whether freedom of expression is more important than privacy, but whether at that point in time "the advantages that the publication of that piece provides, with regard to the freedom of expression and information, outweigh the disadvantages that the publication causes by affecting privacy, reputation or social identity" (Sartor, 2015, p. 73). Such a dynamic approach with regard to the freedom of expression and information in the balance of interests in relation to time, can do justice to the potential changing relevance of information as its accuracy over time is affected by changes to individuals, social contexts, persons of interests, dominant views in society, etc. In this sense, we should not place the core focus on time in its literal, physics, form. Instead, the element of time should be considered in the form of *social time*. With social time, I refer to what we experience as the passing of time due to social transitions, whether it be the transition from a child to an adult, changes in the political climate, lifestyle changes, or simply 'getting older'. The phrasing 'for exercising the right of freedom of expression and information' seems to argue in favour of such a dynamic approach in relation to the passing of time, since 'exercising' suggests an activity in the here and now. Unfortunately, it will not be possible to give strict guidelines as to when there is still an ongoing act of exercising their right to freedom of expression, and when the signifying object is merely the residue of such an act. This is a sliding scale that should be assessed based on the specifics of the case.

Another important factor that matters for the interest that the public may have in certain content, is the person that is the subject of the publication. In European law, public persons receive less protection by privacy rights than the common public, because (at least a part of their actions and life) is a matter of public interest. Although the exact scope of a 'public figure' differs somewhat throughout the European legal rulings, the core of the weight lies in people who hold a public office or play a role in public life, like politicians, public officials, business people and members of particular, generally regulated, professions.[27] With regard to

[26]The other side of the coin is that in some cases the distanced audiences might not be able to (fully) understand the meaning of certain content and are therefore also less of a problem. However, I expect that in these cases, the data subject is less likely to be helped with erasure focused on these distanced audiences and that her interests in erasure lie elsewhere.

[27]Guidelines on the implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and [X]" C - 131/12, p. 13.

people who are less of a public figure, but who do at a certain stage play a role of importance in the public debate, Dommering argues that their private life is likely to legally only be considered a matter of public interest to the extent that it has a direct relation to the public debate itself (Dommering, 2005).[28] While public persons receive less protection of their personal narrative compared to the common public, they still enjoy some protection. Next to the obvious protections like protection from slander and identity fraud, public persons are also granted a certain amount of protection to prevent some of their personal information from being part of their public narrative identity. In *Von Hannover I*[29], and its more nuanced follow-up *Von Hannover II*[30], the European Court granted Caroline von Hannover, despite being a public person, the right to a certain amount of informational privacy. The main point of these cases (although expressed in a stronger version in the first case) is that even in the case of public figures, the (re)telling of the public figure's life story should have relevance for the public.

Lastly, while the privacy interests of the data subject are often positioned as an individual right against the societal value of freedom of expression and information, Lynskey reminds us that we should not forget that the right to privacy and data protection also serve societal objectives (Lynskey, 2015, p. 523). As discussed in section 9.2, the right to privacy protects our autonomy in ways that are important for the (healthy) functioning of a democratic society. Seen from a broader perspective, the eventual erasure of content is therefore not necessarily opposed to the interests of society at large.

The content The last element that I will discuss in the light of the balance of interests, is the role that the nature of the content plays. While this is easily a study in itself, I find it important to at least touch upon the main factors listed by the GDPR, those that have played a role in the (post-)*Google Spain* erasure case law, and factors that are recommended by WP 29 to be taken into account[31].

First of all, with regard to content targeted by an art. 17 GDPR request, it is important to evaluate whether the revealing of that specific personally identifiable information is *necessary* for the purposes of the processing: if and to what degree

[28]The importance of the public status of a data subject may come with a complication: a now common individual may become a public person at some point in the future, resulting in an unforeseen interest of the general public in her history. However, given the ex nunc review of art. 17 GDPR requests, I think that this potential future complication will not play a role in the balance of interests. As the future cannot be foreseen with sufficient likelihood as to who may develop political or similar ambitions, this is a just approach. Else, we would have to retain the information of everyone, just to make sure that we have the information about the very few people that in the future actually become figures of public interest.

[29]ECtHR, 25-06-2004, application no. 59320/00 (*Von Hannover v. Germany*).

[30]ECtHR, 07-02-2012, application no. 40660/08 and 60641/08 (*Von Hannover v. Germany No. 2*).

[31]The factors recommended by WP 29 are part of an advice that specifically sees to search results. However, these factors also seem valuable for the balance of interests in case of content on regular web pages, albeit in a somewhat different weight ratio compared to the balance of interests with regard to search results. I will discuss the balance of interest of search results in section 9.5.1.2.

does the individual needs to be referred to in a personally identifiable manner in order to fulfil goals for the public interest and/or the interests of the controller? If the specific identity of the data subject is not important for the purpose of the processing, the controller may suffice with pseudonimised or anonymised content. For example, in a scientific publicly accessible online publication on the experience of a group of students with social media, it is sufficient (and generally desirable) to use anonymised references to the students in the public report of the research, while the identity of the students who co-operated is locked under restricted access for verification purposes only in case the scientific integrity of the article is questioned.

Secondly, the nature of the content and its sensitivity for the data subject's private life matters.[32] I expect that especially the categories of sensitive information that are given extra protection by art. 9 GDPR, like information referring to health, sexual preferences, race and religion, can carry significant weight in favour of the interests of the data subject. While the processing of these special categories of information is allowed, it is tied to more restrictions than regular personal information. Even if this information is processed according to the restrictions, I can imagine that the sensitive character of the information may affect the balance of the interests more quickly in favour of the data subject than would be the case with more regular information.

Thirdly, it matters whether and to what extent the information is accurate, up to date, and relevant for the general public.[33] The less accurate, relevant for the general public, and up to date the content is, the more it is eligible for erasure under art. 17 GDPR. Closely tied to these criteria, is the question of whether the information is properly framed. For instance, if a personal opinion is framed as verified fact, this may easily mislead the general public. For this reason, an unjust framing of content places more weight in favour of erasure than clear facts or information that is explicitly presented as an opinion.[34]

Fourthly, the responsibility of the data subject with regard to the content needs to be taken into account. This takes shape on two levels: the subject's role in the creation and publication of the content and the subject's role in the content itself. If the data subject is the publisher of the targeted object, her intentions with regard to this publishing matter significantly. If the data subject clearly published particular information in order to make this public, she does not have a strong case with regard to others who publish descendant objects of this content (e.g., copies, remixes, commentaries, hyperlinks, or search results).[35] Additionally, if the content itself refers to actions or events for which the data subject herself is to blame, she has a less strong case for erasure.[36]

Fifthly, it matters to which parts of the data subject's life the content relates.

[32]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §81.

[33]Guidelines on the implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and [X]" C - 131/12, p. 15-18.

[34]Ibid., p. 17.

[35]See e.g., Rechtbank Overijssel, 25-01-2017, ECLI:NL:RBOVE:2017:278.

[36]See e.g., Rechtbank Amsterdam, 07-01-2016, ECLI:NL:RBAMS:2015:9515.

WP 29 argues that information relating to the current working life of the subject should in general receive a certain weight against erasure.[37] Content that relates to the private life of an individual, or a past working life, should thus, in general, place more weight in favour of the data subject's interests.

Sixthly, it matters whether the content refers to an individual as adult or as child. In the GDPR, children are considered to be vulnerable persons (recital 75). The regulator decided that mistakes on the level of information sharing should be easily forgiven in the case of children because they are less aware of the risks and consequences (recital 38). It is important to underline that this protection relates to *content referring to children*. The emphasis of this protection, particularly in relation to art. 17 GDPR, lies on cases in which the child gave consent for the processing (recital 65). As an adult, the data subject can request the erasure of content that relates to her as a child (if the data subject herself is still a child, the request should be made by a guardian). Additionally, if a controller processes the content based on art. 6(1)(f) GDPR, content referring to the subject as a child receives extra protection. Art. 6(1)(f) sees to processing that is necessary for the purposes of the legitimate interests pursued by the controller or by a third party. While this ground makes much processing possible, it does require a balance of interest as the processing is only considered lawful if the interests of the controller are not "overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child". Given the extra protection given to personal content referring to children (although the degree of protection differs per case), the interests of the data subject will in general be given extra weight in cases where the content refers to her when she was a child.

Seventhly, the source of the content is also a factor that needs to be taken into account. Information that originates from professional journalistic sources tends to be given extra weight with regard to the freedom of expression and information.[38]

Lastly, the consequences of the particular content matter. If the information causes prejudice against the data subject, has a disproportional privacy impact on her, or puts her at risk, this is in favour of the erasure of the content.[39]

Balance overview In a world with online personal information, individuals can find themselves, on the one hand, stuck with a narrative that suffers from its flexibility at the hands of others: as these others publish and remix content about the referent, they can (unintendedly) hijack her online narrative identity by realising certain identity options for the referent. On the other hand, some references may gain a persistent longterm presence and can thereby become part of the main plotline of the referent's materialised narrative identity and define her

[37] Guidelines on the implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and [X]" C - 131/12, p. 15.

[38] See e.g., Rechtbank Limburg, 22-03-2018, ECLI:NL:RBLIM:2018:2751.

[39] Guidelines on the implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and [X]" C - 131/12, p. 18.

on a global scale. The individual has an interest in a right to erasure to battle both of these extremes. Meanwhile, the interests of the controller, expresser and general public also need to be taken into account.

Overall on regular web pages, the signifying objects are subject to a relatively strong intentionality of the controller and the audience; the content is commonly the result of an human intentional action by an expresser, and the audience needs to take active and generally directed steps to access the content if it is not mediated by a search engine or a feed (I will discuss these two manners of access in sections 9.4 and 9.5). Due to this strong role of human intentionality with regard to basic websites, I expect that these generally enjoy a relatively strong protection by the right to freedom of expression and information. With regard to regular web pages, I argue that art. 17 GDPR is most suitable for addressing cases where the problem is caused by references that have a persistent presence and thereby become lasting dispositions of the refrent's online narrative identity — especially if this is established at the hand of a strong impact of technological intentionality. Such representations may hamper further development of the individual's narrative identity. However, whether such a representation should be adjusted by means of erasure, depends on various factors.

In table 9.3.1.2 I have listed the various factors that are likely to play a role in the balance of interests, and the side to which they commonly (should) add weight. It is important to note that not all these factors are stable: their impact and weight can change over time. With the passing of time, the relevance of information may reduce to the point where the individual's interests in having the content erased prevail. This was one of the main points in the *Google Spain* case. Additionally, also the factor of time is not a factor that in all cases has a clear effect. While I mainly pointed out that the impact of the passing of time works in the favour of erasure due to the outdated character of content, it is important to acknowledge that the impact of time may also work in another direction. With the passing of time, people are likely to change, and may start to significantly differ from the identifying elements in the contested signifying object. For example, throughout their lifetime, people may change their name, their appearances, addresses, country they live in, profession, etc. In *M.L. and W.W. v. Germany*, the impact of time on the appearances of individuals was taken into account. The court argued that as people are likely to become less recognisable in photographs with the passing of time, the weight of pictorial content in favour of erasure can diminish.[40] However, it is questionable if this argumentation will hold in the years to come: while the diminishing impact of pictorial content with the passing of time may have worked with regard to the human eye, with the increased use of facial recognition software, the recognisability of individuals in pictorial content over time is likely to eventually be upheld — at least with the help of mediating technology. The same likely counts for other identifying elements that may change over time: by combining information from different databases, many relations between individuals and signifying objects may be revealed. While

[40]ECtHR, 28-06-2018, application no. 60798/10 and 65599/10 (*M.L. and W.W. v. Germany*), §115.

some of these technologies are not available to the general online public and certain changes to the individual will reduce her identifiability as referent, with an eye on technological developments, I still placed the passing of time as a factor in favour of erasure.

Weight in favour of:	
retention	**erasure**
adult	child
regular information	sensitive information
public figure	non-public individual
current relevance	outdated
controller intentionality	technological intentionality
audience intentionality	technological intentionality
shared background	no mutual point of reference
correct information	incorrect information
necessary information	redundant information

Of all these factors, especially the respective role of the human and technological intentionality should be taken into account. Tipping the balance of interests in favour of the individual in cases where the technological intentionality has the upper hand in establishing an unwanted materialised narrative identity, would not only benefit the individuals themselves, but it can also be beneficial for the general public. As users become familiar with the affordances of the Web, some of them may become too fearful too publish online if no option exists to distance themselves from a once uttered opinion. A voice that is never encoded into the public debate as a result of the expresser's insecurity or fear of the expression's longterm shelf life and/or a distrust in the framing of the content once it 'runs wild' on the Web, likely entails a bigger loss for the public interest than an occasional shortening of the expression's existence. Art. 17 GDPR can therefore also be seen as an asset, instead of only a threat, in safeguarding the freedom of expression and information in the digital age, because it can mitigate self-censorship that a subject might apply out of fear of decontextualisation or a persistence and salience of the content (Gorzeman & Korenhof, 2016). However, for this, it is vital that the erasure is properly balanced. The form given to erasure can help to support this balance, as I will discuss in the next subsection.

9.3.1.3 Erasure and its effectiveness

If the balance of interests tips in favour of the data subject, art. 17 GDPR can require the erasure of certain content from a web page. In this subsection I will discuss options for various forms of erasure and their ability to address the problems the particular presence of the online content may raise.

Starting with the erasure in the form of complete deletion of content from a particular website. By fully deleting content, many of the issues can certainly be resolved. Simply put: what does not exist anymore, does not give rise to problems. In this form, erasure entails the deletion of existing signifying objects for all audiences. Such complete erasure is a good solution for cases where the content itself is the main issue, for instance, by being inadequate or erroneous.

Complete erasure would also likely be a plausible solution in cases where the wrapping of the signifying object in combination with its content construes a problematic narrative. An example would be the placement of someone's beach holiday picture on a porn website: by changing the surrounding, the objective of the content is changed from an expression of a nice holiday on a travel blog to a symbol of sexual gratification in the porn industry. In these cases, the complete erasure of the (descendant) object is also a proper solution.

Another situation where full erasure may be a feasible application of erasure, is when the retention of the information is no longer necessary for the purposes for which it was originally processed or a compatible purpose. In this case, there is no reason to prolong this retention and the content should be erased. Along the same lines, full erasure seems to be a reasonable application of art. 17 GDPR if the data subject withdraws her consent for the processing of her personal information as consent is (too) easily given online. With a few clicks and little thought, a referent can hand over her personal information to a controller. Art. 17 GDPR can function here as a kind of 'undo button' by allowing users to 'undo' the processing of their personal information when they revoke their (technologically too easily given) consent in accordance with art. 7(3) GDPR. This requires the controller to erase the content (unless, of course, the controller has another legitimate ground to continue the processing).

While full erasure can resolve cases such as the ones mentioned above, it is an unsatisfying solution for many audience segregation issues. When individuals want to be able to play different roles in the same time frame, they do not want their materialised narratives erased, but want to keep different audiences segregated from their different storylines. While erasing a signifying object would put an end to possible audience segregation failures, it would also end the sharing of information with intended audiences. Solving audience segregation issues by fully erasing the content is therefore hardly satisfying. The same goes for cases where the content fulfils a particular public interest, but should not necessarily have a high contemporary presence for the global public on the Web. These are for example cases that involve content in archives that, while being outdated, establishes a certain sameness to an individual's online narrative identity due to the content's high and persistent presence. Full erasure is an extreme measure to address a proportionality problem in the narrative. In such cases, it is therefore worthwhile to consider other variations of erasure.

If we let our imagination run wild, we can think of many variations that may realise a partial form of erasure. Although de Terwangne does not refer to these as forms of erasure, she does offer several interesting suggestions that may help to resolve issues if full erasure is disproportional for the scope of the problem and the

involved interests: anonymisation, restricted access, another form of publicity, the dereferencing/suppression of any links, and the addition of extra information that may help to safeguard the context (de Terwangne, 2014, p. 95). Here, I will discuss some of these suggestions as I think they can be understood as a (metaphorical) form of erasure. I will leave aside the suggestion to add extra information because it is too far-fetched to frame this as a form of erasure and bring it under art. 17 GDPR. However, I will briefly touch upon this in chapter 10.

Starting with the application of erasure in a minimal form; we can look for ways to erase only the personally identifying components in a signifying object. This can be done by for example blotching out faces or identifiable physical traits (tattoos, marks, etc.) on images and videos. In texts we can anonymise individuals, or replace full names with partial names or use pseudonyms. Full anonymisation is difficult to achieve, because the cross-referencing of information may easily reveal the identity of an anonymised individual (see e.g. Sweeney, 2000; De Montjoye *et al.*, 2013). However, some basic anonymisation may be sufficient to render an individual unidentifiable for the majority of the audience. Anonymisation and pseudonymisation cut the most obvious direct links to the individual or at least obscure the presence of the particular reference in her identity. As such, the content will not easily become a part of an individual's online narrative identity (although of course, with hyperlinking and the like, the content can still turn into a persistent part of the narrative). Pseudonymisation is already used to address certain problematic cases with regard to the availability of personal information in online archives. A noteworthy example of this, is a case where a Belgian court ordered an online newspaper archive to replace an individual's name with an 'X' in order to prevent the negative consequences of an easy and longterm availability of the information.[41]

The erasure of personally identifiable aspects of content is desirable if the goal is to remove it from the full online personal narrative altogether. Unfortunately, also such partial erasure of content on websites will not allow us to address audience segregation issues. Anonymisation, pseudonimisation and the blotching out of faces, will not be of any help if we want the content to remain accessible for a particular audience. However, erasure under art. 17 GDPR can also take other forms, like the blocking of access (see section 8.2.6). The blocking of access itself can also take various forms. For instance, an art. 17 GDPR request could see to the blocking of online access to specific content, but not require the erasure of the content in the underlying database. The result of such blocking would be that the information cannot be accessed through the Web. This could be of help in cases where for example the global and easy availability of content is problematic for an individual. If in such cases the database is not directly and fully accessible online, but only offline, or in a restricted web environment, the presence of the specific reference may be sufficiently reduced for the individual to uphold her desired narrative identity. Such applications of erasure could be especially useful for information sources like archives, which have a public value, but may

[41] Cour de cassation de Belgique, 29-04-2016, C.15.0052.F/1.

inadvertently turn their content into a main plot element of the individual's materialised narrative identity and thereby present certain references as lasting dispositions of the subjects character.

Erasure could also be applied in the form of a more focused blocking of online access, by for instance blocking the access for IP addresses originating from particular regions of the world. With this form of erasure, the presence of a certain reference can be reduced for a specific group of users. This could solve certain decontextualisation and audience segregation issues. The blocking of access to a website based on the geographical origin of the users, was the sought after solution in the *Yahoo! vs. LICRA* case.[42] In this case, the French organisation *La Ligue contre le racisme et l'antisémitisme* sought to stop the selling of Nazi memorabilia through the an online auction website of Yahoo! in France by compelling Yahoo! to block access to the auction website for French IP addresses.

Also, erasure in the form of blocking can be applied to prevent indirect access to content. Website controllers can consider taking measures that will reduce or even prevent third parties further disseminating the content or establishing links to it. A good example of this is the adding of robots.txt to a website or parts of its content in order to prevent the indexing of (a particular part of) the content by an online search engine. Such an erasure strategy was used by an online archive that (without any legal obligation at the time) chose in a particular case not to make certain content available through online search engines in order to prevent indirect access (Szekely, 2014, p. 41). By installing such restrictions, the online archive required a certain directed intentionality of users to access the content as the users needed the intention to access the archive in order to reach the content.

While these diverse forms of blocking can address certain issues like that of the salience of a particular reference or its availability for certain territorial regions, none of the above seems likely to be able to address audience segregation issues that occur within a small scope (e.g., if the referent wants to segregate audience members from one particular region). As such, the core of audience segregation issues, i.e., the referent being seen in two different roles by her regular (offline) audiences, can not satisfyingly be resolved by art. 17 GDPR — at least not if the problem lies with the signifying object that the referent would like to share with particular audiences. If a descendant object is the trouble maker, this object can be addressed separately.

If we consider all of the above, we can see that the variations in the processing of digital information allow us to explore various forms of erasure. We should aim to chose a form of erasure that fits best with the various interests that are involved. The tipping of the balance of interests in favour of the data subject does not mean that we should necessarily grant her full erasure; in some, or even many cases, the data subject's problem can be addressed with lighter means than full erasure. If this is the case, we can look for a form of erasure that addresses the issue, while respecting as well as possible the other interests that are involved. The examples

[42] United States Court of Appeals, Ninth Circuit, *Yahoo Inc v. La Ligue Contre Le Racisme et Antisemitisme*, 12-01-2006, No. 01-17424, `https://caselaw.findlaw.com/us-9th-circuit/1144098.html`, last accessed 11-03-2019.

of erasure given in this section are far from exhaustive. These forms of erasure, and their respective benefits and disadvantages, are important to explore in the future for the concrete application of art. 17 GDPR. The identification in the previous chapters of the various elements and their respective role in the coming into existence of problems, provides a foundation for contemplating on such new forms of erasure. The exact boundaries of how erasure can be applied, and who can decide this in various cases, will need to take further shape in future application of the right and case law.

Lastly, it is important to remark that the effectiveness of erasure should be considered with caution, both in its initial application and in its aftermath. Despite best intentions of a controller to reach all third parties that have processed the content (art. 17 GDPR's paragraph 2), a successful application of erasure can easily miss some of the content's descendant objects. Some of the signifying content may remain somewhere online, and can rear its head at unexpected moments. Moreover, the erasure itself can trigger a reaction that may imbue the content with a new presence. This can be caused by people — or even press agencies — who disagree with the erasure and decide to republish the content or a list thereof.[43] This can even happen to the point where it goes viral: it can trigger the Streisand effect (see section 7.3). The application of art. 17 GDPR can thus give rise to a hydra effect, where the information is erased in one location and immediately pops up again somewhere else. Especially in cases of public shaming, individuals feel that invoking a right to erasure may cause another outbreak (Ronson, 2016, p. 203). However, a renewed and even increased presence of the targeted content may also be the result of a data subject who seeks the help of the court to enforce her right to erasure. For example, the *Google Spain* case has drawn a lot of media attention to the data subject of the case as well as to the content that the subject wanted to have erased. Moreover, cases revolving around art. 17 GDPR can become a topic of research (as they have in the present study) and as such trigger a renewed interest in the content and even lead to the publication of new descendant objects.[44]

9.4 Social media

While sharing some of the general characteristics and corresponding issues raised by basic websites, social media's peculiar character raises its own set of problems for data subjects. What sets apart social media from many basic websites is that the controller of social media is not the core content provider. Instead, the users who interact on the platform are. The controller only adds content indirectly by means of the platform architecture that automatically publishes on user actions (see

[43] See the earlier mentioned example of the BBC cases in section 8.2.6

[44] While I tried to restrict the impact of this study for the data subjects as much as possible, I am nevertheless encoding new signifying objects that refer to them. Even with my attempts to minimalise the connection between the discussed case and the individual, the information that I necessarily need to provide combined with some online skills of the reader, will in most cases easily lead to the identification of the data subject.

section 5.4). An additional difference is that the platform is often an online silo that requires an account to access. A big chunk of the social media content is therefore not part of the publicly accessible online narrative. However, within the platform, the narrative in there can quickly build up, spread and become dominantly present on the platform itself, and depending on the platform's permeability, possibly also across platforms or online in general. In their characteristic role on the Web, social media thus tend to configure personal narratives within their own ecology.

Because of the social character of social media, the content on the platform tends to be highly personal and focused on interaction. Social media invite self-broadcasting and are regularly used by users to present themselves, experiment with their identity and associate themselves with particular others, ideas or subcultures (see section 5.4). Combined with their easy editing options and a strong focus on the now, social media have their weight in the shaping of a materialised selfhood — one that might even be over the top; social media give rise to a ongoing realisation of identity choices by inviting a stream of updates on the self and associating oneself with others, and if needed disassociating oneself and editing earlier self-expressions. However, due to the participatory role of the audience on social media, the audience can highly impact the narrative that is told by adding and annotating content and co-shaping its meaning (de Fina, 2016). This results in a personal plotline that is constructed by a stream of micro information composed by users themselves and others, to an often not clearly defined audience (see section 5.6). With its feeds focused on the new and the reacted upon, the construction of this plot is often guided by the value frame and nudging mechanisms of the mediating platform that consistently asks for updates and expressiveness. Taken together, social media give rise to a participatory personal narrative that is hinged on a spectacular constitution of the self.

With its consistent invitation for self-expressions to a difficult to determine audience in an opaque architecture, the main issues raised in the social media's ecology are the result of the encoding of (too) personal information in combination with audience-segregation failures (see section 5.7). Social media are a minefield to navigate social interaction and are prone to trigger spur-of-the-moment actions and judgement errors. Such errors are easily made by the data subjects themselves (like in the Drunken Pirate case). The main issues entail an overrepresention or inadequate image of the self, presented to different audiences in the here and now — a deformed representation of the self shaped in the social medium's triad intentionality of the subject herself, others and the mediating technology. However, as old encoded expressions of the referent remain accessible, and worse, unpredictable in their presence due to the dynamic information flows of social media with their feeds and suggestions, they may resurface, receive attention, and gain a refreshed actuality although they may reflect an outdated self-representation. They may even be imbued with such a presence that they become lasting dispositions in the referent's materialised narrative identity.

In sum, the problems with regard to the manner in which an individual's narrative identity is construed on social media and narrated to audiences, vary from audience segregation failures to the loss of control of an individual over the

construction of her own plotlines.

9.4.1 Applying art. 17 GDPR to social media

While social media share some characteristics with basic websites (the digital affordances of the information, potential ubiquitous access, etc.), they do have a peculiar character because they embody a relation between users and a mediating medium controller. In this section, I will build further upon the conclusions that I have drawn about the functionality of art. 17 GDPR in relation to basic websites and examine how the right functions in relation to social media.

9.4.1.1 Invoking art. 17 GDPR

Given that much of the content on social media is added by users and under their own control (albeit in a joint controller version of situation I), they can erase the content themselves without the need to invoke art. 17 GDPR (for Facebook, this includes comments made by the user on objects posted by others; even the clicking of a 'like' button can be revoked).[45] Moreover, a user can also remove 'tags' referring to her that are placed by others. Nevertheless, in the case of social media, art. 17 GDPR may actually have some use in cases where the control over personal information lies — at least at first glance — in the hands of the data subject herself. Research by Schrems has shown that Facebook retains much of the content that users erased in their database.[46] This content that is retained 'behind the curtains', can be the target of an art. 17 GDPR request. However, discussing this particular application of art. 17 GDPR falls outside the scope of this study because it is not accessible to common users.

The main cases that art. 17 GDPR needs to address on social media, are of personal information that is uploaded and controlled by other users or by the platform. This can be new content (situation V and VI), descendant objects of content originally uploaded by the data subject (situation III), content published by the social media platform itself, like "Paulan is now friends with Maria" (situation VII), or a republication of user content in the feeds (situation VIII).

The first challenge for the data subject is to locate the problematic content. Due to the relatively closed structure of many social media and the wide array of privacy settings, the content may be difficult to locate, or may not even be visible at all to the subject. Moreover, because the subject's overview of what is made accessible by others, as well as which audiences can access what, is clouded at best, she will likely invoke art. 17 GDPR only after she has already experienced some problems as a result of the content and thus knows that it exists and may have some indication where it is.

[45] There may exist social media where comments on publications of others cannot be revoked. This case would then resemble situation II. The corresponding balance of interests will be similar to what has been discussed in subsection 9.3.1.2. However, because discussing all possible variations in detail is too extensive, I will not go into further detail into this hypothetical situation.

[46] Europe versus Facebook. `http://www.europe-v-facebook.org/EN/Data_Pool/data_pool.html`, last accessed 23-11-2018.

Secondly, the data subject needs to identify the controller. While in the early stages of the Web pointing out (not necessarily identifying) the controller was reasonably simple, this is less clearcut in the social media ecology. On these participatory platforms, determining the means and purposes of the information processing has often become a hybrid affair (Koops, 2011, p. 237) — with the exception of situation (V), which entails a publication by the platform itself and thus has a single controller. In their Opinion on online social networking, the WP 29 identifies three possible controllers when it comes to social media: (1) the social media provider, (2) application providers,[47] and (3) users.[48] While identifying the social medium controller should be easy, identifying controlling users may be a challenge as they may use a fake name. To what extent individuals will be able to get identifying information of these users from the social medium controller will depend on national law (I already touched upon this in section 9.3.1.1). Moreover, given the cases of the participatory construction of content, like a thread, it may be difficult to decide for any particular case who the controller is to whom an individual should address her erasure request. However, art. 26(3) GDPR provides significant relief to these issues by allowing the data subject to exercise her right against each of the controllers. With regard to this joint controllership, it is important to note that it does not have to be equally shared: "Different degrees of control may give rise to different degrees of responsibility and liability" [49]. Especially given the power asymmetry between users and social medium controllers, I expect the social medium controller to be attributed a greater responsibility with regard to the information processing that takes place on the platform. Nevertheless, as uploader of the content, the user controller also plays a vital role and the medium controller will need to take her interests into account when assessing an erasure request.

Due to the joint controllership, the subject can invoke her right to erasure against either of the controllers, and depending on the content that she targets, base her request on grounds (a) to (f) of art. 17(1) GDPR: (a) the information is no longer necessary; (b) the subject withdraws consent; (c) the subject objects to the processing; (d) the information is unlawfully processed; (e) the information needs to be erased for legal compliance; and (f) in case of information referring to the data subject as a child. These usable grounds on which the data subject can base her request in case of a social media platform are relatively similar to that of regular web pages. Which ground is most suitable depends on the specifics of the case. Overall, the most promising ground is ground (c), the subject's right to object. This is because I expect that a relevant part of the processing by user controllers that is experienced as problematic by data subjects, is likely to be lawful, still necessary for the purpose of the user controller and not based on

[47] These are the controllers of applications that run on a social media platform, like games. I did not discuss these applications separately in chapter 5, and will not further discuss this specific group of controllers. For the purposes of the present study, their role as controller is comparable to that of the medium controller.

[48] WP 29, Opinion 5/2009 on online social networking.

[49] WP 29, Opinion 1/2010 on the concepts of "controller" and "processor", p. 33.

consent.

As pointed out, there is also a part of the content on social media that is not under joint control. These are the new objects published by the mediating technology, which are only under control of the media controller (situation VII and also VIII[50], to the extent that users have no control over the presentation of these descendant objects in the feeds). An example of situation (VII) is the automatic publication that a user likes a particular group, for example, "Bob likes Kinky Fashion" or "Bob joined Alcoholics Anonymous". If such content relates to the data subject as a child, the most suitable ground for erasure requests in these cases is ground (f), as the request is directed specifically against information society services in relation to their offering of these services. Recital 25 refers to the definition of 'information society service' in Article 1(1) point (b) of Directive (EU) 2015/1535, which states: "'service' means any Information Society service, that is to say, any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services". Social media without doubt tick the boxes of 'at a distance', 'by electronic means' and 'at individual request'. The remuneration box is generally also ticked, even though the users are not paying money for the use of the platform. We can find the inclusion of advertisement revenue models and the like in recital 18 of the eCommerce Directive. Herein it is stated that 'society information services' "in so far as they represent an economic activity, extend to services which are not remunerated by those who receive them". When the content relates to the subject as an adult, the most suitable ground is ground (c): the subject's right to object.[51]

Lastly, it is important to assess whether the content falls within the GDPR's material and territorial scope. Starting with the territorial scope. As many social media controllers have a branch located in the EU and target EU users (see section 8.2.4), the joint controllership construction could place a significant number of the signifying objects on the platform under the control of a controller that falls within the GDPR's territorial scope. However, the material scope, more precisely the household exemption, may place a significant part of the content on social media outside of the GDPR's scope. The core question here is whether the processing of social media users is covered by the household exemption.

In order to fall under the household exemption, a user needs to operate "within a purely personal sphere, contacting people as part of the management of their personal, family or household affairs"[52]. In order to qualify for the household exemption, the information processing needs to be restricted to social and household purposes. Information processing for professional purposes falls outside the scope of the exemption. Additionally, the public needs to be limited

[50]While situation (VIII), the republication of content in the feeds, seems a less likely target for a art. 17 GDPR request given the focus on the now of feeds and the speed with which content drops off the page, it may still be a target of an erasure request.

[51]Due to the opt in opt out character of social media tied to the consent of the data subject to platform policies (see chapter 5), an erasure request based on ground (b), the withdrawing of consent, will likely only work if the data subject seeks to fully end her participation on the platform.

[52]WP 29, Opinion 5/2009 on online social networking, p. 3.

to a *restricted* group of *self-selected* contacts.[53] WP 29 states: "A high number of contacts could be an indication that the household exception does not apply and therefore that the user would be considered a data controller"[54]. Unfortunately, WP 29 does not give an indication of what would qualify as a high number. An extra problem here may be that audience settings can diverge and be complex, raising the question of what kind of settings qualify in order for the household exemption to apply. An example that is worthwhile to take a closer look at in this context, is Facebook's setting where content is made accessible to 'friends of friends'. While the user selects this particular setting and its scope may result in a relatively small audience (e.g., when the user as well as her connections only have a few connections), this audience composition results from the mechanisms of the platform's connective architecture and the user herself did not select the specific contacts that can access the content. In the composition of audiences by means of such settings, we thus see a relatively strong expression of technological intentionality. Because these technologically-driven selections transcend the collection of contacts that are intentionally *self* selected by the user for her household use, I argue that such settings should not qualify for the household exemption — even if they only cover a small number of people.

Despite the limits of the household exemption, I still expect a significant part of social media use to fall within its protective scope. In these cases, social media users will be exempted from the duties of a controller — even when they process information about others.[55] The household exemption can therefore pose a serious challenge for data subjects. An important issue here, is how social media are used. According to the WP 29, "users should only upload pictures or information about other individuals, with the individual's consent"[56], while "SNS also have a duty to advise users regarding the privacy rights of others"[57]. This is a well-intended recommendation, but the reality of social media is that it is a rather common practice to upload content without consent, or often even without the referent being aware of creation or forwarding of the content. This is further complicated as social media are used for communication, which often entails a quick direct interaction, without a previous check about the information that is communicated. Also, the spread of content in a small circle may in itself already be problematic for a data subject. Take for instance the spread of a humiliating photo of someone posted in her family group that contains twenty people. While a selection of twenty contacts likely qualifies as household use, the relatively small circle in which the content is spread is still highly problematic for the data subject. Moreover, the household exemption entails a risk due to the medium's highly networked environment: the

[53] WP 29, Opinion 5/2009 on online social networking, p. 5-6.

[54] Ibid., p. 6.

[55] The household exemption is, however, no get-out-of-jail free card; it does not prevent users from being liable for certain actions in national civil or criminal law (WP 29, Opinion 5/2009 on online social networking, p. 6-7). Moreover, the household exemption is "constrained by the need to guarantee the rights of third parties, particularly with regard to sensitive data" (WP 29, Opinion 5/2009 on online social networking, p. 6).

[56] WP 29, Opinion 5/2009 on online social networking, p. 3.

[57] Ibid., p. 3.

use of content within one 'household', does not mean that the object stays there. Even if an object is spread only *within* small contact circles that would fall under the household exemption, it can hop from 'household' to 'household' by means of contacts that are part of more than one household circle until the content de facto has a massive audience. This is especially a risk due to the high conductivity of social media, that with its mechanisms like share buttons heavily simplifies the spread of information. In theory, the spread can easily become wide; on Facebook, each user is on average only three and a half connections removed from any other user (Edunov *et al.*, 2016). However, this happening under the household exemption is not very likely as hopping from household to household would take time, and might peter out rather than spread very wide. Nevertheless, the protection of social media content use under the household exemption may overall pose a serious problem for the data subject.

It still remains to be determined to what extent people uploading information on social media have to comply with the GDPR (see the article on legal and policy implications of amateur controllers by Helberger & van Hoboken (2010)) or what the household exemption means in cases of joint controllership where one controller falls under the household exemption and the other not. Because WP 29 specifically discussed the scope of the household exemption on social media, I conclude that — at least according to WP 29 — the household exemption can in theory apply to content on social media despite the joint controller structure with the social media controller who cannot make a claim to the household exemption.[58] However, the opinion of WP 29 on social media dates from almost ten years ago (2009), which is a lifetime of development with regard to online applications. I can imagine that the active and problematic processing of personal information by Facebook that recently came to light[59] may be reason to argue that in social media cases in which the content is processed *not only* for household use, the household exemption should not apply — or at least to the extent that the processing transcends household use. However, this is speculation on my behalf. I expect that the Data Protection Authorities, who will supervise the protection offered by the GDPR, will develop more contemporary interpretations of the household exemption alongside the developments in social media. Depending on how its scope develops in the future, the household exemption could severely hamper the effectiveness of art. 17 GDPR with regard to its application on social media as a significant chunk of personal information may hereby be brought out of art. 17 GDPR's reach.

[58] WP 29, Opinion 5/2009 on online social networking, p. 6.

[59] In 2018 it came to light that Facebook improperly shared the data of 87 million users with Cambridge Analytica, a political consultancy agency. See `https://www.bbc.com/news/technology-43649018`, last accessed 25-08-2019. And in 2019, it turned out that Facebook hired third-party transcibers to transcript the audio chats between users that took place on Facebook's Messenger platform. See `https://www.wired.com/story/facebook-voice-transcripts-capital-one-security-news/?verso=true`, last accessed 25-08-2019.

9.4.1.2 Balancing the interests

Assuming that certain content on social media is the legitimate target of an art. 17 request, the most important exception that may stop art. 17 GDPR from being applied, is the freedom of expression and information of social media users. This balance of interests should have roughly the same character as with regard to basic websites, albeit with the main difference that we see a stronger hybrid intentionality in the creation of the content. As social media offer uniform layouts with prefabricated actions, some of the expressive acts entail a high degree of technological intentionality. The most striking of these is the 'like'. By clicking a 'like' button, a user expresses her opinion on an object. However, the specifics and content of this opinion is heavily shaped by the technological architecture: the user merely clicks. In this context, it is interesting to note that a US court ruled that expressing a Facebook 'like' is protected under the freedom of expression. The court stated: "On the most basic level, clicking on the 'like' button literally causes to be published the statement that the User 'likes' something, which is itself a substantive statement"[60]. The court argued that "[t]hat a user may use a single mouse click to produce that message that he likes the page instead of typing the same message with several individual key strokes is of no constitutional significance"[61]. In this case, the US court thus appreciated the intention of the users with the expression of a 'like' equally to the expression of their views with a more elaborate and effort requiring message. However, because the freedom of expression enjoys a broader and more straightforward protection in the US than in the EU (cf. Nieuwenhuis, 2011), I doubt that an EU court would hold the exact same perspective. While I certainly find it defensible to understand a 'like' as an expression that can be defended under the freedom of expression, I am reluctant to treat it as equal to expressions that more strongly express a particular human intentionality. Instead, I opt for a sliding scale in the degree of protection granted to the expressions. The degree of protection would then depend on the respective impact of the technological and human intentionality, while also taking into account the relevant factors in the balance of interests with regard to the freedom of expression and information that were discussed in the previous section regarding basic websites. With regard to 'likes', I doubt that these will be the target of an art. 17 GDPR request because a 'like' is always a strictly defined annotation that itself does not refer to a subject, but only indirectly through its attachment. However, social media entail a high variety of signifying objects, many of them expressing thoughts and feelings of users more elaborately than 'likes'. I therefore expect the freedom of expression and information to protect a significant part of the content on social media.

However, due to the typical joint controller structure on social media, the balance of interests may come with a twist. If a controlling user is exempted from responsibilities by the household exemption or claims protection under the freedom of expression and information, the subject can still try to invoke art. 17 GDPR

[60] 4th U.S. Circuit Court of Appeals, *Bland et al v. Roberts*, No. 12-1671
[61] Ibid.

by requesting the social media provider to erase the object. While it is doubtful that the protection received by the user controller would be overruled as a result of the joint controller structure, the social media controller may choose to enforce the data subjects request anyway because she agrees with the data subject, feels like erasure is the easiest solution, or because she may fear sanctions if she does not comply with the request. Media controllers may be more sensitive to comply with erasure requests because, as Sartor points out, the interest of the medium controller in a particular piece of information is generally lower than that of the publisher, which in turn may result in her choosing to remove content prematurely (Sartor, 2015, p. 92). Here we can find a potential tension between the interests of the medium controller and the user controller. This tension reflects the different values that users and medium controllers attribute to information on social media. As I discussed in section 5.5, information on social media generally has a use value for users, while it has an exchange value for the medium controller. These different value frames can lead to different views and expectations with regard to what should be retained and what erased. However, when a disagreement on this level occurs between a user and a medium controller, medium controllers will have the last word because they control the architecture and decide what the platform policies are. I expect that the general terms and conditions of the media controller may play a pivotal and legally unsatisfying role here; many social media reserve the right to erase content of their users. This illustrates a more general problem with enforcing GDPR rights: most often, subjects are dependent on how controllers act, and while controllers are subject to oversight, it is likely that many controller decisions are not actually subjected to a review by Data Protection Authorities (who cannot oversee everything). However, if a user disagrees with a particular instance of erasure by the social media controller, she can bring the case to court in order to challenge the removal of her content.

9.4.1.3 Erasure and its effectiveness

With regard to erasure of content on social media, we can again think of different forms of erasure that can be applied when the data subject has a rightful claim to erasure. Starting with full erasure. Full erasure of a signifying object could address issues on social media which revolve around problematic content, content that has received problematic annotations or is published in a problematic context. Especially when the prolonged availability of the content itself is the cause of problems, the full erasure of particular signifying objects can address these cases as long as these objects are not spread out of control (see section 9.3.1.3). Additionally, by erasing the reproduction of content in the platform's feeds, the visibility of particular references can be reduced. While this may not be able to prevent the reference's initial presence in the feed, it can prevent a future (re)emergence and a revival of the object's popularity. However, with its mechanisms that focus on the 'now' in a consistent stream of popularity and attention, the application of art. 17 GDPR will likely often be a case of closing the stable door after the horse has bolted.

Moreover, as already pointed out in the section on web pages, fully erasing content does not resolve one of the major issues that result from social media use: audience segregation failures. Partial erasure may be of help here, albeit a little. Similar to basic websites, erasure in relation to social media could take the form of a partial erasure of names, blotting out of faces in images, etc. Moreover, it could entail the erasure of identifying elements typical for social media, like tags. By reducing the identifying elements, objects may drop out of the narrative for particular users because the users do not come across the object, or fail to recognise the referent. As such, partial erasure could reduce the impact of audience segregation failures. However, it does not really address the issue because its main focus lies on addressing the object and not the audiences-access to the object. The original 'Drunken Pirate' case discussed in chapter 1 is therefore an example of a case that art. 17 GDPR cannot resolve.[62]

All in all, the mechanisms of art. 17 GDPR are not well-suited to effectively address some of the main problematic aspects of the participatory narrative that emerges in the social media ecology. Art. 17 GDPR only seems suited to addressing social media cases in which the content in itself is problematic, under the control of others, and is not covered by the household exemption or protected under the freedom of expression and information. And even then, a successful art. 17 GDPR claim may come too late to stop the damage given the particularly high conductive and interactive character of the medium.

9.5 Search engines

Narrating is the act of bringing a story to an audience. As retrieving mediators on the Web, search engines play a pivotal role in bringing stories to audiences as they mediate between audiences and online publishers. They affect the narrative act of the original publisher by directing new, and possible unforeseen and unintended audiences to the signifying objects on her website. However, before a user is redirected to content on another website through the search engine, the search engine offers the searching user a narrative of its own with its search result list and autocompletions. In this section, I argue that, despite the fact that search engines use content produced by others, they are narrators because they produce a new and overarching story for the searching users by bringing together a combination of references in a plot about importance, albeit with a very thin storyline. I will explain this in more detail. Because audiences that access original content through a search result are necessarily first confronted with the search result overview which likely affects their view on the original content, I will keep the main focus in this section on the overarching narrative presented by this overview.

When offering users search results, search engines do not narrate the story exactly as it is authored by the original content providers. Instead, they appro-

[62] An additional reason that this case could not be resolved by art. 17 GDPR is that the content was under the control of the subject herself and therefore outside the working scope of art. 17 GDPR.

priate the content and give their own spin to it. Search engines turn original content into descendant objects in the form of search results and give a kind of commentary on its value by displaying it as a snippet and placing it in a ranking (see section 6.4.3). By selecting, framing, organising, and presenting excerpts of original objects in a ranked collection of search results, the search engine takes on the role of an external authoritative voice that tells an audience what is valuable. In this role, the search engine becomes the narrator of a new story in which it realises a certain emplotment: it configures a set of references originating from multiple narrators into an overarching new narrative and sets the context and audience for the story — although this story's plot is limited to a ranking mainly based on the attention value of the content and/or its source. On the website of the search engine, a narrative is thus construed based on request. Moreover, the search engine may nudge audiences towards particular narrative angles by means of autocomplete.

The implications of the narrative emplotted by the search engine ripple through in the relation between the audiences and the original content: as the search engine turns the search string into the topic of the search results, it can bypass the intentions of the original authors of the publications with regard to the manner in which the story is presented to audiences, as well as the audiences to which it is presented. As the search engine frames the original content in its own narrative plot and directs it towards a particular audience, it can affect the meaning that an audience gives to the original content when this audience follows a search result link. Additionally, the search engine may easily expand the intended audiences of the original sources, not only over time, but also in space — potentially crossing unforeseen contextual or cultural boundaries (see section 6.6).

In particular, the search engine can become the narrator of a 'personal narrative' when a subject's name is used as input for the query. De Mars and O'Callaghan point out that when a search engine does this, it "does not simply 'list'; it constructs online identities of data subjects without their input" (de Mars & O'Callaghan, 2016, p. 227). In this narrative, aspects of original signifying objects are lifted out by zooming in on certain personal references and are weaved into an algorithmically driven plot. By displaying parts of these objects in a ranking, search engines offer users a narrative revolving around attention value with regard to a particular name, where the personal name involved is given a highlighted presence; the plot thickens around the data subject. By combining many references, search engines can even tell 'a life story' of an individual. However, this narrative can be a particularly thin or skewed one. With the priority given to content with a high attention value, the presented personal narrative is quickly boiled down to a few moments of an individual's life or aspects of identity — thereby turning the referent into a 'flat' character (see section 6.6). This effect is strengthened by user behaviour, as search engine users tend to look only at the top results. Especially in the case of individuals with a limited online trail of personal information, the search engine's algorithmically driven narrative can easily establish a particular reference as a core element of the materialised narrative identity by always presenting it as a top result, thereby suggesting it as

an important and possibly even unchanging aspect of the referent's identity. This can reduce the freedom of the data subject to shape her identity. Moreover, the plot can misrepresent the individual if it displays erroneous or decontextualised references. Also, certain references can be given a salient role in the personal narrative, while they, in fact, do not revolve around the individual, or only do so in a marginal manner and have little meaning for the referent herself.

Combined with the dominant use of their services, the narratives constructed by search engines can easily play a defining role in a Web users' interpretation of people. However, due to this crucial role of search engines in Web use, a careful balance of interests is important.

9.5.1 Applying art. 17 GDPR to search engines

Of the situations listed in section 9.3.1, search result cases relate to situation (VIII): a mediating technology publishes a descendant object of the content published in situations (I-VII). Autocompletion cases, on the other hand, are examples of situation (VII): a mediating technology publishes a new object about the subject. With these two functionalities, search engines have a significant impact on online personal narratives. In this subsection I will discuss if, and how, art. 17 GDPR can be applied to traverse some of this impact.

9.5.1.1 Invoking art. 17 GDPR

In the *Google Spain* case, the CJEU argued that search engines determine the purposes and means of processing of the content they present and should therefore be regarded as data controllers.[63] As such, data subjects can invoke art. 17 GDPR with regard to content in search engines if these also fall within the GDPR's territorial scope. Given the relatively broad interpretation of the territorial scope with regard to notions of 'having an establishment in the EU' and the processing of information 'in the context of the activities of this establishment' search engines can easily fall within this scope (see section 8.2.4). However, there is some discussion with regard to the territorial scope of erasure, to which I will return at the end of this section.

Identifying the controller in search engine cases is relatively easy: this is generally a single company that operates the search engine. Also, getting in contact with the controller to invoke art. 17 GDPR should not be too difficult as most search engines have contact information listed somewhere in their help menu. Google Search even has a specific form to file right to erasure requests, albeit specifically for search results of queries that include the data subject's name.[64] While invoking a request to erasure with regard to a search engine should thus be relatively easy, the grounds on which a subject can invoke art. 17 GDPR, is a

[63] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[64] See Google's erasure request form, `https://www.google.com/webmasters/tools/legal-removal-request?complaint_type=rtbf&hl=en&rd=1`, last accessed 15-08-2019.

more complex question that requires some elaboration. I will therefore discuss the various grounds in relation to search engines.

Ground (a) Ground (a) could be a successful ground to invoke art. 17 GDPR against search engines, but only in cases where the processing of information by the search engine is no longer necessary for its purposes. The purpose of the information collection and processing of search engines is to offer users a ranked display of online available information based on a particular query. If the information is no longer available on the source website, the search engine loses its ground to process the information because it no longer offers users an accurate list of the content that is *available* on other pages.[65] When a search engine produces such 'dead' search results, it therefore engages in processing that is not necessary for its purposes. In these cases, a subject should therefore be able to successfully invoke ground (a) against a search engine.

Ground (b) Ground (b), the withdrawal of consent, is likely to be of little use with regard to search engines as the data subjects generally never gave explicit consent to the search engine controller to index their information in the first place. While website controllers can explicitly *not consent* to processing of their content by search engines with the use of robots.txt, this is an opt out instead of an opt in, as I discussed in section 6.3. The use of robots.txt as an expression of the source controller and/or subject's wishes, thus has a dissenting character. The default opt in situation will not suffice to count as consent under the GDPR because here consent requires an *affirmative* action of the subject (art. 4(11) GDPR). As consent is never explicitly given in an affirmative action, it cannot be withdrawn. This does raise the question on which legal grounds the search engine is processing the information. I will briefly get back to this when I discuss ground (d) and (e).

Ground (c) Ground (c), the right to object, seems one of the most fruitful grounds to invoke a right to erasure against a search engine, whether it be against a search result, or against an autocompletion. In the *Google Spain* case, the right to object was the main ground on which the data subject requested a right to erasure. However, it is important to note that the *Google Spain* case was trialled under the DPD, with the DPD version of the right to object (art. 14 DPD). While the right to object in the GDPR (art. 21 GDPR) is somewhat expanded compared to article 14 DPD, the core of the right is roughly the same. Ground (c), as discussed in section 8.2.8, needs to be invoked by an individual in relation to her particular situation. In the case of processing by search engines, the subject could make a case that particular search results that are displayed, present a decontextualised, outdated or erroneous reflection of who she is (now) to the searching user. Search engines would need to demonstrate compelling legitimate grounds to be allowed to continue the

[65] This is also the sole case where Avocate General Jääskinen argued in favour of erasure by search engines in his opinion for the *Google Spain* case, Opinion of Advocate General Jääskinen, 25-06-2013.

processing of this specific narrative plot, i.e. the display of these particular search results. An example of a case where there were compelling legitimate grounds to continue the processing of certain search results, was a Dutch case in which the data subject once was a news item because he was accused of using academic titles that he did not rightly hold. In this case, the court argued that because the subject was a public figure and desired to work in education, internet users had a legitimate interest in being able to access this content through a search engine.[66]

Ground (d) and (e) Grounds (d), which sees to personal information that is unlawfully processed, and (e) where the information has to be erased for compliance with a legal obligation, could be potentially successful grounds to invoke against search engines.

To start with, the legal obligation grounds come with a challenge, especially given the global scope of the Web; different countries worldwide may uphold different laws on what constitutes lawful processing. This is a challenge for multinational corporations that run online search engines, but seems to be a reasonable cost for doing business globally. However, as the laws can vary per country and discussing different national laws lies outside the scope of this study, I will focus here on lawful processing according to the GDPR, and more specifically, on ground (d).

In order to get more grip on the limits of lawful processing by search engines, it is important to consider the legal grounds within the GDPR that allow search engines to process personal information. Search engines are most likely to process personal information on the ground listed in art. 6(1)(f) GDPR: "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child". This ground explicitly requires a balance of interests that takes the interests of the individual into account, thereby clearly opening the door for art. 17 GDPR requests.[67]

The search engine's grounds on which it processes information are more problematic with regard to the special categories of personal information (art. 9 GDPR). Kulk and Zuiderveen Borgesius question the legal base of search engines to index[68] web pages that contain sensitive personal information (Kulk & Borgesius, 2014). Art. 9 GDPR prohibits the processing of sensitive personal information, unless the controller bases her processing on the grounds listed in art. 9(2) GDPR. For search engines, the most likely grounds are 9(2)(a) GDPR "the data subject

[66] Rechtbank Midden-Nederland, 20-11-2018, ECLI:NL:RBMNE:2018:5594, §4.18.

[67] While a search engine may attempt to base its processing on ground 6(1)(e) GDPR, which covers processing for the performance of a task carried out in the public interest, this claim will likely not hold in courts in the EU due to the search engines' profit models.

[68] As Google Search scrapes, stores, and organises the content of websites without having a contract with the website controller that marks the search engine as processor, nor received an assignment of her to index her content, the search engine operator determines the purposes and means of the processing and thereby performs the role of controller with regard to the content that the search engine indexes.

has given explicit consent to the processing of those personal data for one or more specified purposes", and (e) the personal information is "manifestly made public by the data subject".

Starting with ground 9(2)(a), processing of sensitive information based on explicit consent from the data subject. Search engines do not have such consent from the subject when they index web pages and publish part of its content as search result. However, this is part of a bigger issue, and solving it is rather complicated. As Kulk and Zuiderveen Borgesius rightly remark, it will be impossible for search engines to get consent from all the individuals mentioned on the websites that they index.[69] The lack of consent would render a significant part of the information processing by search engines illegal (Zuiderveen Borgesius, 2016, p. 223).

However, consent may not be needed in all cases. A search engine can also invoke the exception listed in art. 9(2)(e) GDPR as a ground to process special categories of personal information. This exception states that the processing of special categories of personal information is allowed when the processing concerns personal information that is *manifestly made public by the data subject herself*. This may be a viable ground for the processing of special categories of personal information with regard to the information that is publicly available on the Web. The question here is how 'manifestly' should be understood. While the details still need to take shape in case law, I think that in general for the term 'manifestly' to apply, it is sufficient that a data subject explicitly uses a particular medium that comes with a reasonable expectation of public access to the content on this medium. However, the question following this, is whether Google Search checks whether the content is indeed uploaded by the data subject herself. I suspect that a considerable part of the sensitive information on the Web is uploaded by others (this is for example what happened in the Lindqvist case discussed earlier). Given the fact that in the *Google Spain* case, Google Search argued that "search engines process all the information available on the internet without effecting a selection between personal data and other information"[70], it is safe to assume that the search engine also does not check (at least when indexing) whether the uploader is the data subject as this would require the search engine to differentiate between personal and other information. Hence, while there may be grounds for Google search to process special categories of personal information, it is questionable whether Google Search is able to recognise which information it can legally process before it actually processes it. Sensitive information made public by others than the data subject are a critical problem for search engines. The result is that if a

[69] I do see a role here for the website controller, as she is responsible for the disclosure and dissemination of the information. A web page controller could for instance tag the parts of the website that contains sensitive information with robots.txt to prevent its indexing. Although website controllers do have a certain responsibility here as they are the controller and in that role responsible for the disclosure of the information, placing the full burden of the indexing actions of search engines with the web page controller, seems unfair. Unfortunately, I currently see no other options if we want to maintain the current functionality of search engines.

[70] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §22.

search engine indeed lacks a legal ground to process the information, a subject can successfully invoke art. 17 GDPR on ground (d) against the search engine.

Ground(f) Ground (f), which sees to the information collection in relation to the offering of information society services to a child based on consent, has little significance for the erasure of search results or autocompletions. In recital 18 of the Directive on Electronic Commerce (2000/31/EC) search engines are categorised as information society services: "Information society services span a wide range of economic activities which take place online; these activities can (...) extend to services which are not remunerated by those who receive them, such as (...) those providing tools allowing for search, access and retrieval". Ground (f) can thus be invoked against search engines, but it only targets the information that search engines collect about their own users. This is not the content that is presented *as* search result or *as* autocompletion (but it does shape the search results and autocompletions for a particular user). Additionally, there is the question of whether the manner in which search engines collect information about their users by means of cookies qualifies as consent under the GDPR due to the power imbalance between the search engine and its users (cf. Zuiderveen Borgesius *et al.*, 2017). Ground (f) is thus of little use for subjects who are confronted with a problematic narrative as result of the presentation of their personal information to other users by search engines.

In sum In sum, the most promising ground to invoke art. 17 GDPR against a search engine is ground (c), on the basis of the subject's right to object. The core of this success lies in the relatively weak grounds that search engines have to process personal information. Where art. 6(1)(f) GDPR allows search engines to process personal information without consent, it immediately curbs this freedom to the point where the interests, rights and freedoms of the data subject become preponderant. With regard to special categories of personal information uploaded by others, or by the individual herself but clearly for a restricted audience, the data subject could also invoke her right to erasure on ground (d).

9.5.1.2 Balancing the interests

In this subsection I will discuss the balance of interests with regard to content presented to users by search engines. As discussed in subsection 9.3.1.2, the balance of interests is embedded in several parts of art. 17 GDPR. I expect that most of the balance of interests in relation to search engines will take place in connection to ground (c), the right to object, which I touched upon in the previous subsection, as well as in relation to exception (a) that sees to the right to freedom of expression and information. However, due to the particular character of search engines, the weighing of the interests may lead to different results than it would in cases that concern regular websites. In the *Google Spain* case, the CJEU argued that "the outcome of the weighing of the interests at issue (...) may differ according to whether the processing carried out by the operator of a search engine or that carried

out by the publisher of the web page is at issue, given that, first, the legitimate interests justifying the processing may be different and, second, the consequences of the processing for the data subject, and in particular for his private life, are not necessarily the same"[71]. What is interesting here, is that the CJEU considers search engines to have a more severe impact on an individual's privacy than the publication of content on a regular website.[72] Given my analysis in chapter 6, I agree with the CJEU on this and endorse the corresponding difference in weight that the CJEU attributes to the various interests between content presented by search engines and content on regular web pages. In search engine cases, the stakeholders are the data subject, the general public, the search engine controller, and the controller of the original source to which the search result links. Because the information processing of search engines like Google Search is based on art. 6(1)(f) GDPR, they necessarily need to perform a balance of interests between the stakeholders.[73]

In this subsection, I will split the discussion of the balance of interests into two parts. First, I will discuss the balance of interests in relation to search results. Next, I will discuss how I think this balancing act differs in autocompletion cases.

Search results As I argued in the previous sections, when balancing the interests of the controller(s), the data subject, and the general public with regard to online accessible information, we would do well to assess these interests in relation to the impact of the mediating technology: the impact of the technological intentionality should serve as a complementary evaluative ground in weighting the different interests. So too in search engine cases. In the case of search engines, the publication of content (search results) takes place in response to a query and entails a strong technological intentionality on two levels: (1) the original content is molded by the search engine into a descendant object that tends to differ highly from the original object; and (2), as various results are placed together in response to a query, the search engine resemiotises the original content by absorbing (snippets of) it in a new algorithmic driven narrative that revolves around the importance of the content for the query. By doing so, the mediating technology more strongly affects the presented narrative than it does in the case on basic websites where usually human publishers (whether they be the controller or not) create the content and context. To balance the interests in a manner that does justice to the impact of the mediating technology as well as to the different stakeholders, it is important to have a closer look at this role of the technology in the formation of the hybrid intentionality of search engines in relation to the interests of the controller to disseminate information as well as that of the original publisher, and the interest of the public to gather information. I will start by discussing the search engine operator's right to freedom of expression, and will follow with a discussion of the interests of the original content controller. Last, I

[71] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §86 and §86.

[72] Ibid., §86 and §87.

[73] See e.g., Rechtbank Amsterdam, 19-07-2018, ECLI:NL:RBAMS:2018:8606.

will discuss the general public's right to information.

Starting with the interests of the search engine controller. Search engine controllers process content originally shared by others and offer this in a new narrative wrapping established by the selection and ranking algorithms of the search engine. The search engine controller does not produce a particular narrative herself, but instead lays out the rules for a technological narrator, while the content of the narrative is brought forward by the mediating technology. One of the fundamental questions in the balance of interests in this regard is therefore to what extent this processing is covered by the right to freedom of expression and information. Van Hoboken, who did extensive research on this topic, argues that when presenting search results, search engines "combine a passive (conduit/access) and active (editorial/selective) role" (van Hoboken, 2012, p. 209). He further argues that this selective role, even though performed by algorithms, implies the making of editorial decisions, and therefore deserves protection under the right to freedom of expression (van Hoboken, 2012, p. 209). However, the protection that a search engine like Google Search can receive from the right to freedom of expression is curbed by its business model as the "right to freedom of expression does not protect commercial and noncommercial communications to the same degree" (van Hoboken, 2012, p. 173). Van Hoboken therefore concludes that search engines that function as advertisement platform, like Google Search, have a weaker claim to protection than they would have had if their funding was comparable to, for instance, public libraries (van Hoboken, 2012, p. 172). This view was confirmed by the CJEU in the *Google Spain* case. The economic interests of Google Search were given specific consideration by the court. The court argued that, while a publisher of a website can publish information for journalistic purposes, a search engine often reproduces this information for other purposes: Google Search does "not merely give access to content hosted on the indexed websites, but takes advantage of that activity and includes, in turn for payment, advertising associated with the internet users' search terms"[74]. With this, the CJEU separates the intention of the search engine operator in the production of search results from the intention of the publishers of the original content.[75] Although the CJEU does not directly discuss the freedom of expression of search engines in the *Google Spain* case (a lacuna in their argumentation, rightly criticised by e.g., Kulk & Borgesius (2014)), from this position that the CJEU holds in the case we can derive at least some indication of how to apply the right to freedom of expression to search engines. The separation of the intentions of the original publisher and of the republishing search engine (operator) is an important and, as I will argue, justified step in the balance of interests. However, I suggest to deepen this split a bit further given the functioning of the mediating technology.

To the degree that the aggregation of a search result list is algorithmically performed and based on websites that are automatically crawled, I argue that the content displayed by search engines can only indirectly be understood as the

[74]CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §43.
[75]Ibid., §85.

expression of thoughts, feelings or interests of the search engine controller (leaving aside those results that are manually selected and given priority). Only to the degree where the search engine controller prioritises certain sources, languages and the like in the design of the search engine's algorithms, do the results reflect the thoughts of the controller. This differentiates the publication of search results, next to the purpose of the processing, from many of the publications of content on basic websites or social media (aside from feeds) and viral publishers: in these cases the creation of the signifying object results generally from an action of a human publisher that expresses a direct (not necessarily deep or well thought through) intentional relation regarding that particular content (please note, that I am focusing here on the publisher of the content, which is not necessarily the controller, see for example situations IV and VI). We can thus assume that in most of these cases, there is a direct relation between the publication of that particular signifying object and the *particular* views, thoughts, feelings or interests of the publisher. Contrarily, in the case of search results, the publication of the content — its reframing in a technologically calculated new narrative —, leans for the majority on technological intentionality: the intentionality of the search engine controller is not directed towards the specific content of a search result. I therefore take search results to be of a lesser value for the expresser (which in this case is the search engine controller) than publications that have stronger roots in human intentionality (thus in most cases the original publication). While I attribute some freedom of expression to search engines, their 'expressing value' should weigh less as there is no direct link between the human intentionality of the controller and the specific presentation and framing of particular content. An appeal to the freedom of expression by a search engine controller, should therefore be given a weaker position in the balance of interests, not (only) because of the commercial character of search engine service delivery, but (also) because of the strong technological mediation in the presentation of the narrative.

While the interests of the search engine controller may not carry significant weight in the balance of interests, this is different when it comes to the controller of the original content. Because a significant portion of the Web traffic goes through search engines, many publishers have a strong interest in being accessible by means of search engines. Making content more difficult to find can even be regarded as a restriction of the freedom of expression of a publisher (see e.g., van Hoboken, 2012; Kulk & Zuiderveen Borgesius, 2015). This is the point where art. 17 GDPR should be balanced with the accessibility of information. Accessibility plays an important role in the freedom of information, because information that is encoded but completely unaccessible, may as well be considered as not being there at all. However, respecting the interests of the original publisher in making her content accessible through search engines may be a problem in certain cases. With the strong protection given to sensitive information and the position of search engines as independent controllers under the GDPR, search engines will have almost no legitimate ground on which they can lawfully process sensitive information like political opinions (art. 9 GDPR) or criminal convictions (art. 10 GDPR) shared by others than the data subject herself (Zuiderveen Borgesius,

2016; Kulk & Zuiderveen Borgesius, 2017). This is a dilemma because the access to this information can be pivotal for the freedom of expression and information of the original publisher and the general public. This is a difficult situation, and legal scholars are still discussing how to best resolve this (see e.g., Zwenne *et al.*, 2015; Zuiderveen Borgesius, 2016). A possible solution is that the Data Protection Authority would grant search engines an exemption to process, under certain circumstances, sensitive information without consent (Zwenne *et al.*, 2015, p. 17). However, how this will finally be dealt with is not clear yet. For the time being, and in the light of the importance of the freedom of expression and information of the original controller and the general public, I will assume that this will be resolved, and that search engines will be allowed under certain circumstances to process sensitive information that is shared by others than the subject herself.

While the original publisher can highly welcome, and even depend on, the mediation of her content by search engines, it is important to underline that in connecting audiences with her content, the search engine impresses its own intentionality on this connection: the content is zoomed-in, reframed and assimilated in the search engine's own ranked narrative. The original publisher thus has little say in how, in what context, and for which audiences her content is made discoverable. Search engines may even reveal information to audiences unintended by the publisher (see for example the LGBTQ case discussed in section 9.3.1.2). I therefore propose to approach the interests of the original publishers in a contextualised manner that is in line with their intentions while taking into account any unintended or unnecessary impact of the search engine's intentionality. I suggest looking at the role that the query plays in relation to the content — and thus in relation to the original publisher's intentions. Contextualising the interests of the original publisher will likely be most valuable in relation to name searches. For example, if the name is one of many, used as an example, or plays a minor role in the original signifying object, we may assume that the intentionality of the original publisher is not directed at that particular referent. In these cases, a reduced findability based on specifically a name query has a less significant impact on the freedom of expression of the publisher compared to when, for example, the name refers to a referent that plays a key role in the content.[76] In such cases, the interests of the original publisher should carry ideally less weight *in relation to that particular query* than it would in relation to certain other queries, like a name query that sees to an individual who is attributed a key role in the publication.

Next to the interests of the search engine controller and of the original publisher, the freedom of individuals to gather information by means of search engines also needs to be taken into account. Courts have recognised that search engines play a

[76]The importance of the data subject for the content of a particular signifying object will need to be checked per case. For example, a website on "Friends of Mort" will list many names, but that does not mean that the content is not significantly about Mort. Additionally, if the name is hidden in a large set of names, it is less likely to carry much weight in the search engine's algorithm, and therefore is less likely to show up in the top-10 or 20 search results. On the other hand, if it is a not so common name, it may indeed end up high in the top. There are thus many variations possible, all with their own specifics for the balance of interests.

pivotal role in helping individuals to access information.[77] However, this interest needs to be balanced. For this balance, the nature of the content and the role of the data subject in societal life should be taken into account.[78] For this part of the weighing of interests, I would like to refer the reader back to section 9.3.1.2. Also, in case law we can find another factor that is relevant to take into account in the balance of interests specifically in relation to search engine cases: the character of the source of the content seen in relation to the availability of the information in other sources. In a Dutch case, a surgeon wanted to have search results removed that pointed to an unofficial 'blacklist' that referred to the fact that she was disciplined by the medical board. Here, the court took into account that the information about the surgeon's error and the subsequent disciplinary action could easily be found in the official disciplinary registration for medical professions.[79] Partially due to this availability of the information in another easy to access online source, the court considered the prominent role attributed to the unofficial 'blacklist' by the search engine as unnecessary and excessive.

Moreover, it is relevant to take the user input into account. The public gathers information by means of a query and thereby expresses a certain intentionality by entering a particular string (I will leave aside autocomplete for now, and discuss that in the next subsection). This needs to be taken into account in the balance of interests. The more personally focused the query, the higher its implications on the narrative identity: a personal name search may reveal a set of references revolving around a particular referent. However, in a less or non-personalised query, the focus lies elsewhere. Given the differences in impact and overview, I argue that in order to erase results that are returned in response to a less or even non-personalised query, higher standards should be met. An example of a less personalised query is a search string like 'top executive [company] lives in container'. The string 'top executive [company] container' was part of a Dutch court case in which a top executive wanted to have search results erased that referred to a conflict he had with a constructor and due to which he and his family were forced to live in a container in their backyard.[80] While the words on their own do not directly identify a particular individual, the combination of them, especially in Dutch, does. This would be less the case with for example the search string 'living in a container', which does in itself not refer to an individual. Yet, in some cases a data subject may want to see even search results in response to such non-personalised search strings removed (generally in addition to erasure of the results in more personalised queries, like in the case of the Dutch top executive). At the moment it is unclear whether and when art. 17 GDPR can successfully be used to address search results in relation to non-personalised search strings. So far, I have not come across cases in which an erasure request with regard to

[77] See e.g., CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*); Rechtbank Midden-Nederland, 20-11-2018, ECLI:NL:RBMNE:2018:5594.

[78] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §81.

[79] Rechtbank Amsterdam, 19-07-2018, ECLI:NL:RBAMS:2018:8606, §4.16

[80] Rechtbank Amsterdam, 13-02-2015, ECLI:NL:RBAMS:2015:716.

personal search results in response to a specific non-personalised search string was granted.[81] However, there is nothing in art. 17 GDPR itself that would rule out such requests or would justify turning them down immediately without balancing the interests.

While the user input expresses a certain intentionality, it is relevant to consider how the search engine aims to change this into output, a search result list, that is of interest to the user in order to fulfil her information gathering wishes. As the search engine 'interprets' the user search string in its own frame of reference, it produces results according to its own logic (see section 6.4.1). Due to search engines' algorithms that are inclined towards the production of an a-chronistic overview based on a 'spectacular scheme', they can easily present the old and the marginal as top result or decontextualise information. As such, search engines may inadvertently present something as characterising for the referent, while it poorly reflects her own sense of self. The question that this raises is whether such a mischaracterised portrayal of a data subject is of interest to the information gathering individual. The emphasis lies here not on the separate results, but on the overview and the suggested importance in ranking the various results. The question underlying this, is on what grounds we should decide what is more or less important about a particular individual for the general public. In the *Google Spain* case, we can find some interesting factors listed by the CJEU that may be of help to decide if particular search results should be used to represent a data subject in search engines. The CJEU states that individuals have a right to have search results removed if they are "inaccurate (...) inadequate, irrelevant or excessive in relation to the purposes of the processing, that they are not kept up to date, or that they are kept for longer than is necessary unless they are required to be kept for historical, statistical or scientific purposes"[82]. In this phrasing, these elements also seem to encompass situations where information has become excessive or inadequate when, for instance, a minor element of a bigger signifying object is disproportionally highlighted or decontextualised.

One of these factors I would like to highlight: time (the content needs to be kept up to date and only retained as long as needed for its purpose[83]). Search engines can appropriate the informational history of a referent, while she herself has no control with regard to how she is represented by the search engine. If a certain source with old, possibly even outdated content, has an authoritative status, the top search results may easily contain references to this content. This is what likely played a role in the *Google Spain* case: despite the fact that the content referred to an event in the relatively distant past, the reference was shown as one of the top search results. The consequence is that it can become difficult for individuals to successfully distance themselves from past views and actions in the eyes of

[81] For this, I consulted the Dutch case law collection on `www.rechtspraak.nl`, articles by various legal scholars who discussed art. 17 GDPR cases, as well as reports of journalists on 'right to be forgotten' cases.

[82] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*), §92.

[83] Ibid., §93.

those who use the search engine. Sartor's view on the shifting balance of interests with regard to the passing of time is therefore especially relevant with regard to search engines (Sartor, 2015). As with the passing of time, the 'newsworthiness' and freshness of information tends to decline, and the content loses some of its relevance and meaning in the context of the public's right to information (Korenhof *et al.*, 2015, p. 191). Sartor therefore argues that "[w]hile there may be a strong public interest in name-based access to a page containing fresh news about a person, this interest is likely to decrease as time goes by and be outweighed by the person's interest that the information is not accessed in this way" (Sartor, 2015, p. 96). How to shift the balance of interests with the passing of time in concrete search engine cases is a research topic on its own. However, an example of how to practically employ the factor of time is offered by a Dutch court. In a recent case, the court used the possibility for a former convict to receive a 'statement on behaviour' (verklaring omtrent het gedrag) as a measure for the passing of sufficient time.[84] These statements are issued by the state and give citizens clearance to fulfil particular professions. Depending on the profession, citizens should not have been involved in criminal offences the last four or ten years in order to receive such a statement. The court thus argued that the possibility to receive such a clearance statement indicated that sufficient time has passed to justify the erasure of search results referring to a particular criminal offence.

In sum, in the process of performing a search — from input to output — the mediating technology of the search engine presses a rather strong technological intentionality on the framing of the narrative that it presents in the search result overview, on the audiences that it connects to the original content, as well as on the context in which it connects them. In this process, search engines can easily construct a particular plot in the search result overview that characterises the referent based on outdated information or decontexualises or mischaracterises more contemporary representations of her. If we focus on the intentions of the diverse stakeholders as we balance their interests, we can try to identify the cases in which the search engine affects a subject's materialised narrative identity in ways beyond direct human intentionality. Especially in these cases, there is a reasonable ground to strongly consider erasure of the targeted search results. In those cases where there is clear direct human intentionality involved in the display of a particular search result in relation to the used search string — i.e., the intentionality of the search engine controller (unlikely), that of the original controller (more likely), and/or that the searching public (more likely, but less in case of an autocompletion), or better, of more parties —, the balance of interests should resemble more closely that of basic websites. However, in these cases compared to basic websites, extra weight should be added to the interests of the data subject due to the search engine's more severe impact on an individual's identity construction than basic websites (see the beginning of section 9.5.1.2). The above discussed factors can help to fine-tune the attribution of weight given to the interests of the diverse stakeholders.

[84] Rechtbank Amsterdam, 15-02-2018, ECLI:NL:RBAMS:2018:1644.

Autocompletions As discussed in chapter 6, autocompletions can reveal information, certain relations, and can even convey false messages to the searching user. The impact of autocompletions on the online personal narrative is significant, as they lead the user towards particular plotlines. Autocompletions have been the target of dispute in court cases[85] and gave rise to a legal discussion that shares many similarities with the discussion surrounding the question of whether a search operator can be regarded as the controller of search results (cf. Karapapa & Borghi, 2015). Here, the *Google Spain* case[86], again, plays a key role. The CJEU's argumentation in this case that the search engine operator should be considered a controller of the search results, can also be applied to autocompletions (Karapapa & Borghi, 2015, p. 282): the search engine operator determines the purposes and means in which the personal information is processed, and is in the case of autocomplete even the sole holder of the content because is not collected from elsewhere. As such, it is safe to assume that autocompletions can be successfully targeted with an art. 17 GDPR request.

However, there are significant challenges with regard to assessing autocompletion erasure requests because autocompletions can affect individuals with the same name (Karapapa & Borghi, 2015, p. 269-270); users may confuse the person they search for with the person used in the autocomplete. For example, let us take the name 'Peter Murphy'. This name occurs in Wikipedia's list of names that are included in human name disambiguation pages.[87] These are pages where multiple people with the same name are listed, in order to make sure that the content is matched to the intended referent. Peter Murphy is inter alia an artist, a footballer, a singer, a politician and a businessman. An (hypothetical) autocompletion in a search engine, like 'Peter Murphy bankrupt', may easily reflect on multiple Peter Murphys. While for the bankrupt Peter Murphy 'Peter Murphy bankruptcy' may be a justified autocompletion, this autocompletion is problematic with regard to the other Peter Murphys who are not bankrupt.[88] How should controllers and courts deal with erasure requests filed by a Peter Murphy who has never been bankrupt while a bankrupt Peter Murphy does exist? I have no one-solution-fits-all answer for this. Such erasure requests need to be judged per case. Factors that are relevant to consider are whether the name is common, the size of the risk that the name may be attributed to a wrong referent, the scope of the problems for the wrong referent, as well as the interest of the general public in having access to that particular autocompletion.

[85] An example is a Dutch case in which the plaintiff requested the erasure of an autocompletion that tied the name of a famous Dutch crime reporter to the subject's name. Gerechtshof Amsterdam, 31-03-2015, ECLI:NL:GHAMS:2015:1123. In this case, the court turned down the plaintiff's request for erasure of the autocompletion.

[86] CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).

[87] See "Pages that link to 'Template:Human name disambiguation'", `https://en.wikipedia.org/wiki/Special:WhatLinksHere/Template:Human_name_disambiguation`, last accessed 20-08-2019.

[88] On a side note: this can work both ways. For example, Donald Trump (not the president of the US) may have all kinds of embarrassing facts about him online, but the chance of a search engine suggesting those in autocomplete is negligible due to his more famous namesake.

So far, the balance of interests performed in the autocompletion court cases tended to have either a strong focus on the source of the autocompletion (users, the search engine operator), or on the impact that it has on the recipients, the searching users (Karapapa & Borghi, 2015, p. 277). As argued throughout this study, a focus only on either of these two, or even on both, overlooks the possible impact — and intentionality — of the technology itself. To balance the interests of the different parties we should therefore take the respective role of the human agents at both sides of the technology into account, as well as the (hybrid) intentionality of the mediating technology. Taking this into account, I argue that the distribution of weight in balance of interests in the case of autocompletions should have a somewhat different character than it should have with regard to search results, this because human intentionality plays a different role here.

Human intentionality plays a significant role in the creation of autocompletions, because the suggestions are to a great degree based on searches of users. Moreover, contrary to search results, autocompletions are not descendant signifying objects. However, if we compare the human intentionality in the creation of autocompletions to the intentionality of the publishers whose content is displayed in the search results, we see that they differ at a fundamental level: while the content published in search results is generally meant by the original publisher to reach at least a certain audience, the autocompletions display a human intentionality that was never aimed at *expressing* a certain view to others (except, maybe, for those looking to game the system). Instead, users use the search string as input for their own information retrieval. Autocompletions are therefore not expressions of users, but consist of search behaviour that is technologically reshaped and commodified into a search functionality. However, despite the fact that autocompletions are based on information that individuals did not intend to use in order to express themselves, the repeated use of a particular search string is an indication (a) that many of the public are interested in this notion, and (b) that the notion (or the combination of words, whether it is true or not) is widespread and hence somewhat resembling common knowledge. This gives autocompletions a certain weight for the public interest. I therefore argue that the lack of intention of users to express themselves in an autocompletion balances out against the public interest in these completions. This leaves the impact on the individual as the main factor in tipping the balance to either side.

9.5.1.3 Erasure and its effectiveness

I take the current application of erasure with regard to search engines and the acceptance thereof in courts — the delisting of a search result as response to a particular query — to be a confirmation that erasure in relation to search results does not have to entail the erasure of the content from the search engine's database (see section 8.2.6). Applying erasure only in the form of delisting the result in response to particular queries has the advantage that the URL can be returned as search result for other queries. This may be a preferred outcome, even for the data subject. Take for example, again, the BBC BLLCKS cases discussed in

chapter 6. In these cases the subjects co-operated with BBC to raise awareness for testicular cancer. However, they did not want these interviews to be displayed as a search result following a query on their name. Given that these subjects co-operated with a series to raise awareness about testicular cancer, I assume that they wanted people with questions about testicular cancer to be able to access the content. Because search engines play a pivotal role in online information access, I expect that these subjects also would agree to the retrieval of this content by means of search engines with a query focused on testicular cancer and the like. Also, because it is far less likely that people searching for testicular cancer know the subject, in contrast to people searching for their name. These queries therefore have (far) less effect on the subjects' narrative identity as presented to the audiences they personally engage with. The delisting of a search result in response to a data subject's name is therefore sufficient to address problematic returns in those cases where the search engine decontextualises and highlights a particular personal reference and displays a problematic narrative revolving around a name. However, some cases may exist in which a more thorough form of erasure, like the delisting of search results also for certain non-personalised queries, or even the complete deletion of the content from the search engine's database, is needed. If only particular queries are considered to be a problem, delisting the results in response to more queries can be sufficient to address the issue. However, if the content itself is considered to be problematic by the data subject, irrespective of the query with which it is retrieved, the erasure should be focused on the content itself. In case the subject has tried to have this content erased at its origin, and failed, or if she cannot reach the original controller of the content (both of these options would address the issue in a more thorough manner), the subject can try to restrict the object's accessibility by objecting against the processing of the content by the search engine and needs to make a case to request its deletion from the search engine's database.

Next to full erasure and delisting, we can also consider other ways to realise the reduction of the presence of a particular reference in a search engine. If the salience of a particular result that is outdated is the main problem, it is not necessarily needed to fully delist the result to resolve the problem. In these cases, we can also consider downranking the result, thereby removing it from the top results — the main plotline — and giving it a less prominent presence in the overall narrative. Especially given the attention that users generally give to top results, downranking particular search results may suffice to address a part of the issues (although downranking may be technically far more complex than the delisting of a particular object and for this reason not feasible). The suggestion to downrank or reorder search results is put forward by inter alia Stuart (2013), as well as de Mars & O'Callaghan (2016) who with this argue for the introduction of "more nuanced means of addressing privacy concerns" (de Mars & O'Callaghan, 2016). By downranking or reordering search results, information can be presented in "more contextually appropriate ways" (de Mars & O'Callaghan, 2016). Whether the downranking of search results can be considered as a valid form of erasure under art. 17 GDPR is unclear — in time, hopefully case law will tell. Given the

highly nuanced potential of downranking to resolve search result cases, it would be beneficial for art. 17 GDPR's problem solving capability if this is the case.

Moreover, we can consider the territorial scope of the implementation of erasure. Search engines like Google Search offer their services in applications focused on particular domains. They profile their users and redirect them towards the domain that fits the geographic origin of the user's IP address. In law and in literature it has been a lively discussion whether search results should be erased from view only for particular domains of the search application (e.g., .nl, .be, .fr), whether the result should be erased for all domains, or whether the result should only be blocked for users with an IP in particular countries (geoblocking). In some cases, local erasure or geoblocking could be sufficient to address the issue. An example of such a case, briefly discussed in section 6.4.3, is the case of the autopsy photos of a US citizen on a Cambodian website. These photos were considered problematic when they were displayed as search result to a US audience. Making these results inaccessible only to the US based audience, would likely be sufficient to address this case. What is important to take into account with regard to the scope of the erasure, is that in all likelihood the most significant impact for a subject takes place in her local domain: this is where she lives most of her onli*f*e. However, this is also why, in case of a name search the removal on further distanced domains may make more sense: distanced audiences tend to have less to do with the referent and therefore tend to have less of a public interest in the person. Correspondingly, it also matters where the referred to event took place. Other nations have less to do with national and local affairs, and have often a poorer view of the legal and cultural context of events that took place in another nation. As such, they will likely have a more limited understanding of the narrative of the event. A case where cultural differences between countries played a role is the Feldmar case discussed by Mayer-Schönberger in his book *Delete* (Mayer-Schönberger, 2009). Feldmar is a psychologist who experimented with LSD and wrote an article about it from a scientific perspective.[89] When he wanted to visit the US, the border control looked Feldmar up in a search engine and came across the article. Based on this find, Feldmar was refused entrance to the US. Behaviour that might be accepted and even legal in one country, can be a problem in other countries. This can highly complicate our 'onlives', because it is difficult to foresee all the countries we may want or need to go in the future, or how future laws and cultural views may develop globally. The value of information, and the likelihood of users correctly understanding the context of information will generally decline the further the user is removed from the country of origin. These types of problems are arguments in favour of a global scope for the application of erasure.

The territorial scope of the erasure of search results has been a point of discussion. One the one hand, there are those that argue that erasure should only apply to EU based search queries[90], and on the other hand, there are those

[89] Andrew Feldmar, "Entheogens and Psychotherapy", *Janus Head*, 4, 54-67, 2001. `http://www.janushead.org/4-1/feldmar.cfm`, last accessed 01-08-2018.

[90] See for example, Daphne Keller, "Don't Force Google to Export Other Countries' Laws", *The New York Times*, 2018. `https://www.nytimes.com/2018/09/10/opinion/google-right-`

who argue in favour of a global scope[91]. In France, the scope was the point of dispute in a court case between Google and the CNIL (Commission nationale de l'informatique et des libertés, the French Data Protection Authority). The French court referred the case to the CJEU. On 24 september 2019, the CJEU ruled that when a search engine operator grants an erasure request, she is not required to delist the targeted search result on versions of the search engine outside of EU Member States.[92] However, Member States can decide to order a search engine to delist the results globally.[93] Restricting the delisting scope to the EU may leave certain cases unresolved and may leave original publishers in some cases with a relatively all or nothing choice. To take up again the example of the interview of a data subject with a Dutch LGBTQ magazine, it seems highly unlikely that it is the intention of the original publisher that her interviewees are at risk when travelling outside the EU as a result of the easy accessibility of her publication by means of a search engine. However, the original publisher also likely has an interest in being easily found through the search engine by many EU audiences. If the delisting of a search result is restricted to EU versions of the search engine, the controller will need to choose between opening up her content for global access through the search engine (possibly only with the exception of EU access) or to make the content unavailable through search. In cases where the scope of deslisting is restricted to the EU, the burden of restricting the audiences of their content to EU-citizens is placed with the original publisher.

There is a last point that is important to discuss in relation to erasure in search engines: the practical side of erasure. Search engines may need to deal with a massive number of erasure requests (this seems to be the case for at least Google Search, which received 785.489 requests targeting 3.040.510 URLs between May 2014 and January 2019[94]). Controllers who have to deal with such massive numbers of erasure requests need to come to a quick procedure for dealing with these. The result is that they may need to automate the procedure to some extent. Automation necessarily comes with some risks on the level of false positives and false negatives. Additionally, people may seek to exploit the false positives and realise unjustified erasure of content. Typical enough, this brings us back to the issue of a too strong impact of the technological intentionality in the shaping of our information flows. A possible large quantity of art. 17 GDPR requests at the address of a particular controller, can thus interfere with a sufficiently careful balanced application of the right to erasure by requiring the implementation of a strong technological intentionality on the flip side.

Taking everything into account, I conclude that interfering with the narrative presented by search engines can correct the view produced of an individual in a

`forgotten.html`, last accessed 23-11-2018.

[91]For example, WP 29 argues in favour of a global scope in the Guidelines on the implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and [X]" C - 131/12, p. 9.

[92]CJEU, 24-09-2019, C-507/17, ECLI:EU:C:2019:772 (*Google v. CNIL*), §74.

[93]Ibid., §72.

[94]Google Transparency Report, `https://transparencyreport.google.com/eu-privacy/overview?hl=en`, last accessed 18-03-2019.

search engine and thereby effectively address the problems she may experience. Especially given the pivotal role of search engines, applying erasure to search results can be an effective means to address issues with an individual's online narrative identity on a wider scale. While applying erasure in search engine cases can be a very viable way to address problems with the online narrative, it is important to critically assess the balance of interests and choose the form of erasure with care. Erasure in search engine cases is not without risk given the dependency of the general public on search engines. A part of this risk may be mitigated by the fact that despite erasure of the search results (whatever form of erasure is applied), the original signifying object remains available on the source website and users who spend enough time and effort may be able to locate it. However, with regard to small and relatively unknown websites, we need to assume that the content is lost to the general public (which also can be a good thing). I therefore argue that full erasure of the content from the search engine's database should only be applied if there are fundamental problems with the content itself. Given the pivotal role of search engines in online traffic, it is preferable to delist, downrank or geo-block search results instead of fully deleting the content from the search engine's database; this way the content can still be reached through the search engine, but with restricted queries. In the case of problematic autocompletions, the problem is likely to lie with the content of the autocompletion. In this case, full erasure will be of use, and maybe in some cases, geo-blocking or anonymisation.

9.6 Viral outbreak

A viral event has a strong impact on an online personal narrative: a single reference becomes so strongly present that it easily turns into the main plotline of the referent's complete materialised narrative identity. The other signifying objects are likely to be sucked up within this plot created by the viral reference. The viral plot may even establish a leading narrative that caricatures the referent to the extent that she may become represented as a longterm symbolic character. Especially the combination of a viral reference attached to an individual's name and the mechanisms of search engines and social media heavily impact the online narrative identity of individuals (see chapter 7).

Moreover, this impact may not be restricted to the real referent of the signifying object. When a young woman posted a photo showing her screaming and putting up her middle finger next to a grave with the sign "Silence and Respect", the photo went viral and she became target of a massive shaming campaign. As a result, a huge number of signifying objects referring to the incident and mentioning the woman by name, were all over the Web:

> *"That five seconds of her life is her entire Internet presence?" I said. Farukh nodded. "And it's not just this [X X]. Anyone who has that name has the same problem. There are sixty [X Xs] in the U.S. (...) and they're all being defined by that one photograph"* (Ronson, 2016, p. 264).

Reconfiguring a narrative shaped by a viral outbreak requires a fundamental change in the overall emplotment. As I will discuss in this section, addressing the viral impact on a narrative plot is difficult with art. 17 GDPR. However, it is important to note that the virality of a particular reference is not static. A viral outbreak goes through the stages of its initial outbreak, followed by a period of decay, and finally remains lingering in its afterlife (see chapter 7). As such, with the passing of time, the impact of the viral reference on the materialised narrative identity will diminish. Reconfiguring the materialised narrative is therefore, next to the application of art. 17 GDPR or other regulatory approaches, a matter of letting time do its work. However, there always remains the risk of a revival.

9.6.1 Applying art. 17 GDPR to a viral outbreak

Of the situations listed in section 9.3.1, viral cases potentially cover all situations due to the fact that they are often spread across the Web and its diverse applications. As such, the application of art. 17 GDPR in viral cases is likely to run across many, if not all, of the difficulties, as well as the possibilities, that I have identified in the previous sections. Taking the previous subsections into account, I will examine if, and how, art. 17 GDPR can be applied to address (some of) the impact of a viral outbreak.

9.6.1.1 Invoking art. 17 GDPR

In case of a viral outbreak, the data subject will generally quickly become aware that personal information is being processed in a manner that negatively affects her online narrative. However, due to the spread and scope of this processing, she will likely have difficulty locating all the relevant signifying objects and their corresponding controllers. Although art. 17(2) GDPR should be of help here by requiring the controller to take reasonable steps to inform other controllers that process the targeted information that the data subject has requested the erasure thereof, I expect this to have a limited effect with regard to a broad and spontaneous information spread. For the information controller, it may be equally difficult to trace all the third parties who processed the information (see section 8.2.6). Moreover, with a viral spread, it can be difficult to discern which descendant object resulted from which previous controller. Controllers may therefore have a difficult time figuring out the scope of their responsibility.

In the end, I expect that the data subject of a viral outbreak will need to invoke her right to erasure against myriad controllers. In the worst case scenario, the data subject will find herself confronted with all the challenges that I identified in the previous sections, multiplied with the number of controllers that she has to address. This is a burdensome task, and in the case of a large viral outbreak, an impossible one. The subject dependency thus makes it difficult to effectively use art. 17 GDPR for large quantities of signifying objects.

9.6.1.2 Balancing the interests

The balance of interests in relation to a viral outbreak is likely to encompass all of the above discussed balances of interests, with an extra magnitude on the level of the public interest and the freedom of expression. What is typical for virality, is that there is an overlap between the expresser and the public: what was the public, becomes the next expresser.

Because the content is shared massively by the general public to one another, a certain public weight is attributed to it; the content becomes part of Web culture and is treated as a public good (see chapter 7). Despite the fact that the affordances of digital online content play a significant role in the spurt of the viral outbreak, the decision to spread and remix the content lies literally for the majority in the hands of users; users share and manipulate the content. As such, human intentionality is an important component of a viral outbreak — but in many cases, it is a 'thin' intentionality. A significant part of the spreading of content online is achieved with a quick copy paste, or even a semiautomated action like the clicking of a share button, which can be performed with little effort and time for reflective thinking. This is different in the case of remixed content, where people express themselves with creative exploitations of the content (i.e., the many descendant objects of Technoviking or the Star Wars Kid) or provide the content with commentaries. This content ties in more strongly to the freedom of expression. We see this position with regard to viral content recognised — albeit not in those words — by the German court in the case of Technoviking. Although the original publisher was ordered to take the content offline, the descendant objects created by some other publishers in the form of remixes found legal protection because they could appeal to their 'artistic value'.[95] This is especially interesting as the original Technoviking video — the one ordered offline — was, in fact, an art project.

In order for a proper balance of interests in the case of viral content, it is worthwhile to determine to what extent the expresser expresses her idea, thought or feeling with the content. One of the factors we should take into account in this, is to what extent the (re)publication results from an action and effort on behalf of the user. In the case of solely clicking a share button, we can attribute much of the sharing action to the intentionality of the website architecture. This changes, however, when the user adds a personal comment. As in the sections previously discussed, the stronger the publication of the content seems to reflect the ideas, thoughts, and feelings of the expresser, the stronger the content should receive protection under the freedom of expression.

Because a part of the viral cases results from strong emotions (e.g., the Dog Poop Girl), it can be argued that these cases should enjoy a relatively strong freedom of expression. However, I do think that the value of certain expressions online for the expresser, need to be assessed with care and we should hold back with too easily accepting an expression as *necessary* for venting thoughts and feelings. As pointed out in section 4.3.3, online users generally experience a disinhibition effect and feel more free to express extreme thoughts and emotions — thoughts

[95] Landgericht Berlin, 30-05-2013, Nr. 27 O 632/12.

and emotions that they would not display in the same manner in offline settings. The question we should focus on per case (and this may be an almost impossible task for those who need to apply art. 17 GDPR) is: is this expression at the time of the art. 17 GDPR request necessary (proportionality factor) for exercising (temporality factor) the freedom of expression? I can imagine that after the viral peak and the initial emotions are vented, the necessity of the ongoing retention of the content quickly declines, and may reach the point where it does not outweigh the interests of the data subject.

Next to the freedom of expression, it is also important to consider the right to gather information with regard to viral content. While the mass sharing suggests a public interest in gathering the information, I argue that this is less the case than the mass attention initially may suggest. In chapter 7, I connected to the work of Varis and Blommaert, who argue that the core of virality does not lie in the meaning of the content, but instead in its effect: the sharing of content serves more as a social action than as a sharing of information (Varis & Blommaert, 2015, p. 41). The mere virality of particular content is therefore not a convincing marker for the value of this content.

Additionally, if we consider viral content in the light of some of the points discussed in section 9.3.1.2, we see that some content elements that are typical for many viral cases place weight in favour of the data subject. Viral content is often simple, features non-famous individuals, and is made by amateurs (see e.g., Shifman, 2013; Jiang *et al.*, 2014). Due to the simplicity and often mainly entertaining character of the content, the content will generally have less relevance for the general public than political speech and the like. The more limited relevance for the general public places an equal limited weight in favour of retention. Additionally, viral data subjects are generally not people who fulfil a public role in society and it is questionable whether they can be marked as public persons. While they may have become a person of public interest due to the public interest generated by the virality itself, such a self-fulfilling interest loop seems a dissatisfactory ground to mark someone as a public person. If the data subject cannot be regarded as a public person and the content does not relate to her working life, this places weight in favour of the interests of the data subject. Moreover, many viral objects are not published by the referent themselves, and the publisher tends to be an amateur. Also both these elements place weight in favour of erasure (see section 9.3.1.2).

However, some of the points discussed in section 9.3.1.2 are more likely to place weight in favour of the retention of viral content. In cases like Technoviking and Dog Poop Girl, the data subject did have a significant responsibility in the course of action that is displayed by content that went viral: they acted in a certain manner in public space and it is their particular behaviour that is at the heart of the viral reference. Their responsibility in the particular event in public space can place weight in favour of retention.[96] Another element that can place weight in favour of retention, is when official information sources like newspapers reproduce

[96]See e.g., Rechtbank Amsterdam, 07-01-2016, ECLI:NL:RBAMS:2015:9515.

the viral content.

Because viral cases are very diverse, I cannot provide a detailed balance of interests here. However, as the value for the expresser to disseminate the information and for the public to gather the information lies for a significant part in the social value of the action, and not in the value of the information itself, it is in many viral cases questionable if (all) the corresponding signifying objects should enjoy a strong protection under the right to freedom of expression and information. Overall, given the strong disproportional manner in which viral content affects the materialised narrative identity of the data subject, I argue that there is a significant weight in favour of erasure of the content — at least at some locations — so that data subject can free herself from an unwanted strongly emplotted narrative established by a viral reference. I will discuss this in the next section.

9.6.1.3 Erasure and its effectiveness

In cases of a viral outbreak, the problem lies in the online presence of many signifying objects that share a certain reference. The strong presence of a particular reference can easily establish a particular piece of information as the central caricatured storyline of an individual's materialised narrative identity. The main goal therefore is to thin out the presence of this particular reference, by either erasing the signifying objects, or by reducing their identifiable character.

While erasure (whether it be fully or partial) can in theory address the problems caused by a viral outbreak, the subject driven character of art. 17 GDPR is an obstacle for the right's effectiveness in practice. Even a dedicated NGO or a reputation managing company acting on the subject's behalf will likely have a hard time filing the necessary erasure requests to reduce a reference's presence in case of a wide viral spread.[97] The ability of art. 17 GDPR to control a viral reference therefore depends on its spread; as long as it is contained in a single or few platforms under the control of one or a few controllers, like on a social media platform, it seems doable for the data subject to invoke her right against each of the controllers. However, due to the local character of such outbreaks, there is also often a direct link between the data subject and the users viewing the content (for example, they are classmates, colleagues and/or friends). In these cases, erasure of the content will not be able to undo the initial damage: the others have already seen the content and will likely remember it. What erasure could do, is prevent the still small outbreak from becoming a nationwide or global epidemic. For this, time is of the essence: to prevent a further outbreak, the content needs to be erased as soon as possible. Given the mechanisms of art. 17 GDPR, it is questionable whether the right is suitable for such quick application.

While art. 17 GDPR in general is an unfit tool to fully address a viral reference, it may still be used in a manner that can reduce the impact of the outbreak on a

[97] Hence, reputation managing companies tend to take up other strategies like overruling the presence of the problematic reference with other references (see e.g. Ronson, 2016, p. 263-274).

data subject's materialised narrative identity. The first is a local solution, and the second a gatekeeper solution. I will discuss these consecutively.

Firstly, what erasure could do, is to remove (some of) the most prominent viral objects from the direct environment of the referent when the outbreak is in its decay phase. Lessening the number of reminders to the viral event reduces the memory triggers and the potential confrontations with the caricatured view of the referent. As such, erasure may help a data subject to move past a viral event and the particular caricature that is pressed on her, and allow her to shape her own identity again in relative freedom. However, this will only be a local reduction in the online environment of the subject, and will in many cases have little impact on the presence of the reference in the overall Web.

Secondly, by targeting search results or autocompletions of search engines with art. 17 GDPR, the data subject could make a noteworthy dent in the prominent presence of a viral reference. What plays an important role here, is that many viral outbreaks become known under a nickname (e.g., Technoviking, Dog Poop Girl, Star Wars Kid) but not under the person's real name. If we look for example at the Technoviking case, and let us say that his real name is 'Sam Vimes', most people who see the Technoviking reference will not understand this as a narrative of Sam Vimes, but of a character known as 'Technoviking'. The result is that Sam Vimes' narrative identity only suffers from this on the points where Sam Vimes' audience associates him with the Technoviking, which will happen either based on his name or appearance. As Sam Vimes' appearance is likely to change over time, the main linking factor will be his name (at least as long as there is no facial recognition software applied in search engines). Here, art. 17 GDPR as a right to delist may come in handy, because when a search engine is prevented from showing a reference to Technoviking in the search results or by autocompleting Technoviking when people type in 'Sam Vimes', the risk that Sam Vimes' audience will associate him with the Technoviking is significantly lowered. As such, art. 17 GDPR can serve as a partial solution to even the most difficult of the web's information problems, although perhaps surprisingly, in its guise as a right to delist.

9.7 Reconfiguring the narrative with erasure

In this chapter, I have proposed to approach art. 17 GDPR as an identity-related right that can work as a 'counter technique'. With an eye on its capacity as a counter technique, I have investigated the problem-solving capabilities of art. 17 GDPR for the cases examined in this study. In this closing section, I will walk through what I have found to be the right's most important assets, limitations, use principles, and values.

A flexible solution to a group of problems As I have shown in the problem analysis chapters, there is not one problem. Instead, there is a group of problems that have a certain family resemblance in the fact that they influence the online materialised narrative identity beyond the wishes and/or expectations of the

subject. The 'long-lasting memory' of the Web, the kickstarter of the wish for a 'right to be forgotten', is just one of the issues, and not even the most pressing one. A more important genetic trait in this family of problems is the Web's connectivity in combination with the affordances of digital objects: content is easily created, edited, and spread beyond expectations. Users navigate in this realm and may easily engage in what later turns out to be too easy publishing (or, at the other extreme, some may become so fearful of the affordances of the Web that they prefer to refrain from online publishing at all).

The problems for individuals raised by the presentation of online personal information to Web users occur in the sphere of a hybrid intentionality where human agents as well as the mediating technology can be the key factor. While humans who process information online are always expressing a hybrid intentionality, their decisions play an important role in this processing. Online, people are encoding new content, as well as copying, editing, and remixing existing content. Such creation of new signifying objects and descendant objects is often simplified by online service industries that offer simplified options for the encoding of information with the use of WYSIWYG website generators like Wordpress, and social media like Instagram, Facebook and Twitter. In these cases, the encoding and dissemination of content is industrialised to a greater or lesser degree. Despite the role of the mediating machinery and its often limited transparency and control options, users are the ones that make the final choice in the encoding and dissemination of this content. Human encoded content that itself is experienced as problematic by the referent (these range from singular encodings to a viral outburst), is therefore a problem that for an important part can be attributed to human intentionality.

However, overall, the mediating technology seems to be the more prominent factor in the arising of many of the discussed issues. Given this key role of the mediating technology, understanding art. 17 GDPR as being able to function as a 'counter technique' is a fruitful complementary understanding of the right. In this guise, the right helps individuals to counter undesired effects on their public narratives that result from the online processing of their personal information. Especially the affordances of digital information that allow the creation of descendant objects in the form of copies, remixes, search results, and the like, can establish a certain salience and/or decontextualisation that may turn even relatively unproblematic content into a problem for the subject.

The forecasting case law of the *Google Spain* case paved the way for a relatively open interpretation of erasure. The relatively open and broad scope of erasure, as well as the right's capability to address descendant objects while leaving the original object in place, fits well with this goal of the right. Without the capability to address descendant objects while leaving the original object in place, art. 17 GDPR would lose much of its potency to address the problems identified in chapters 4 to 7. By addressing particular descendant objects, the right to erasure can reduce, and even solve, a part of the problems, especially in those cases where the content in its original context is not experienced as problematic for an individual's narrative identity, like in some of the BBC cases (see section 6.3).

Meanwhile, the open interpretation of erasure allows us to look for an application of erasure that is proportional to the interests of others, and that also alleviates the problems for the subject. Erasure could entail a full or partial erasure of the object itself, an erasure of an object from view for (particular) users, or by implementing certain manners of processing that reduce the presence of a particular reference. Ausloos, who did extensive research on art. 17 GDPR, came to a similar conclusion when he states: "One could think of the fifty shades of erasure as a spectrum, with free and unrestrained processing on one side and full erasure at the other end. Rather than oscillating between the two, solutions will generally lie in shifting the needle along this spectrum so as to find the most balanced solution" (Ausloos, 2018, p. 349). However, in order to come to the most balanced solution, I argue that in cases where the problems occur as a result of a too strong technological intentionality, we should fine-tune the form of erasure specifically *to the impact of a particular technological mediation* on the produced narrative.

Lastly, while art. 17 GDPR acknowledges the passing of time as a relevant factor for erasure, its functionality is not limited to erasure in relation to the passing of time. If we look at the different grounds on which art. 17 GDPR can be invoked, especially when based on the right to object, we can think of myriad situations where the right can justifiably be invoked. With its broad application range, art. 17 GDPR can be used to address at least a part of the problems identified in the present study in a proportional manner.

Addressing search results The biggest asset of art. 17 GDPR is its ability to address search results. Of all the technologies I examined, search engines have the biggest impact on the presentation of our materialised narrative identity. As centralised mass access and service points, these third party bulk processors of online content tend to become pivotal gatekeepers in the online information flows. Their architecture expresses a particular power structure of information exchange. Their scope, scale, and detail, combined with their gateway function not only impacts the presence of (certain) references, but can also configure a larger narrative on the basis of signifying objects associated with a personal name. The combination of their pivotal role as gatekeeper with the presentation of a particular plot based on importance and popularity, establishes them as narrators with an authoritative voice. They can construct a zoomed-in and decontextualised highlighted presence of a particular name for even the most marginal original context and thereby establish a high qualitative presence for certain references. Time in their narrative is not displayed sequentially, but is mixed up in the ranking created by diverse algorithms, in which popularity and the authority of the source play a key role. The popular becomes the main plotline of the narrative. From their pivotal position, search engines tell searching users who you are and what is important about you. Seen from this perspective it is hardly surprising that people are not thrilled to be defined by, for example, their once cancerous testicles or their opinion on a video game. In this context it is interesting to note that Reding in her 2012 speech already placed the focus on "unlimited search and memory capacity"

as the big impact factors of the Web.[98] While the unlimited memory capacity turned out not to be one of the biggest problems, the search capacity is.

However, next to — and partially due to — their heavy impact on the narrative identity, search engines are also the most successful point to apply art. 17 GDPR in order to address problematic narratives on the Web in total. In guiding users towards content, search engines prioritise certain references over others. With this, they can increase or decrease the chance that a specific user comes in contact with a certain reference. They thus have a powerful influence on the formation of audiences for particular content. This influence can be used to reduce the presence of certain online references. The strong position and gatekeeping role of search engines imbues them with the power to address a fair share of the issues, or at least reduce the issue in their most intense form. The erasure of search results can even help to mitigate some of problematic impact of a viral information flow on its subject.

Online consent Art. 17 GDPR can be a valuable instrument with regard to a particular subgroup of the problems, namely cases where the content originates from the subject and she has initially given consent for this processing. Online, consent is easily given, probably too easily. The technological affordances of online interaction allow users to easily, naively, drunkenly, sleepily, or in a moment of stupidity, hand over their personal information with a few quick clicks to a controller — and generally without reading what they are agreeing to. I think, except for extremely privacy sensitive users, we all have been there at one of more points in our onli*f*e. When revoking her consent in accordance with art. 7(3) GDPR, a user can subsequently invoke art. 17 GDPR to 'undo' the processing of their personal information based on their technologically too easily given consent by requiring the controller to erase the content (unless, of course, the controller has another legitimate ground to continue the processing). As such, art. 17 GDPR functions as a kind of 'undo'-button that can be activated in corollary to art. 7(3) GDPR.

Consent is also relevant to consider with regard to another technological context: the Web as digital information network. Online, information can be easily copied, republished, linked to by others, as well as indexed by search engines. However, when an individual participates with an online publication, e.g., an interview in a newspaper, she does not necessarily give consent for the indexing thereof by search engines or the further spread of the information by third parties republishing the content. This shows the possible discrepancy between the manners in which users reflect on consent and the heavily networked information practices of the Web. For example, the BBC 'Bollocks' cases indicate that the individuals interviewed, did not take into account that the interviews would be indexed and displayed by search engines as a search result to their name, as they requested the removal of the search results rather quickly after the publication

[98] Viviane Reding, SPEECH/12/26, *The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age*`http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm`, last accessed 4-11-2018.

of the interviews. I therefore have the impression that when users give consent, they often do not fully take the technological character of the Web into account as a heavily networked information realm where their information easily moves into contexts other than the one that they consented to. While these third parties may have a legitimate ground to process the content, their processing does express a clash between the expectations of a participating individual and the networked practices of the Web. As such, these kinds of cases show how important it is that art. 17 GDPR allows individuals to use the right to erasure to target descendant objects of the original object. With this, they can address issues like decontexualisation and the salience of particular signifying objects that followed, for them likely unexpectedly, from their consent.

Allowing individuals some retroactive control with regard to content to which they consented or played an active role in its creation, not only benefits the individuals themselves, but can also benefit the public interest when the erasure is properly balanced. While I do share the general concerns with regard to the potential negative impact of art. 17 GDPR on the freedom of expression and information, I argue that art. 17 GDPR can also be seen as an asset in safeguarding the freedom of expression and information in the digital age, because it can mitigate self-censorship that a subject might apply out of fear of decontextualisation or a persistence and salience of the content (Gorzeman & Korenhof, 2016). A voice that is never encoded into the public debate as a result of the expresser's insecurity or fear of the expression's longterm shelf life and/or a distrust in the framing of the content once it 'runs wild' on the Web, likely entails a bigger loss for the public interest than an occasional shortening of the retention period or an obscuring of the identifiability of the author of these expressions. Viewing art. 17 GDPR as a technique to counter the impact of technology, thus allows us to understand art. 17 GDPR as not necessarily opposed to the freedom of expression and information.

Best equipped for single persistent objects in the public sphere Despite the significant flexibility of art. 17 GDPR, its mechanisms are not suitable or not ideal to address all the problems that I touched upon in this study. For instance, art. 17 GDPR is unsuitable for cases where the individual herself expresses herself to the wrong audience. Additionally, the time that art. 17 GDPR may need to be effectuated, may be too long to address some cases. Sometimes a quick removal is necessary, like in the case of revenge porn.[99] Due to its subject-dependent invoking mechanisms, art. 17 GDPR feels somewhat sluggish for the highly dynamic Web where everyone can copy, upload and disseminate content in mere seconds. Especially with regard to social media, with their high focus on the now, the subject will probably already have experienced problematic consequences of the content before she can invoke art. 17 GDPR. Moreover, the functionality of art. 17 GDPR for social media is further hampered (at least in certain cases) by

[99]See Rejo Zenger, "Sommige dingen wil je écht snel offline", *Bits of Freedom*, 2018. `https://www.bof.nl/2018/06/25/sommige-dingen-wil-je-echt-snel-offline/`, last accessed 01-08-2018.

the household exemption. While the household exemption in theory should work well together with art. 17 GDPR because art. 17 GDPR ideally addresses the (semi)publicly available or organisation controlled materialised narrative identity of individuals, it does potentially place a significant use of online technology outside of the right's functional scope. The risk of the household exemption, given the connective character of the Web, and especially the high connectivity of social media users, lies in the fact that the 'household' character of online information can be rather volatile. A signifying object that is initially only visible to twenty people, can with a few clicks become visible to thousands of users. However, the household exemption is an important exemption to protect individual autonomy with regard to the use of a more personalised tertiary memory; individuals may process all sorts of personal information about others for personally relevant reasons. For instance, someone may keep a diary where she reflects on her difficult her experiences with others, or someone may send an email about an ex-lover to a close friend in order to ask for advice. The risk that some of this information may spill outside the household scope is one that is well worth the interests that it protects, but this does likely leave some problems unattended to — at least by art. 17 GDPR.

Moreover, there is a limit to what the individual can do to correct her narrative identity. While the right gives individuals a certain degree of control, it is precisely this individual control, the subject-driven character of the right, that turns art. 17 GDPR into a burdensome tool to address a viral reference. Even if the subject calls in the assistance of a dedicated NGO or hires a reputation managing company, addressing a wide viral spread is difficult. This is unfortunate, because the impact of a viral reference on the subject's life is likely high. However, this does not mean that all is lost in viral cases. Because virality is often a mix of diverse media, the problem can be reduced by, for example, delisting search results referring to viral content, or by removing the content from the most prominent sources. Moreover, in viral cases, the passing of time actually may demonstrate some of its distancing character. As the viral content moves into its afterlife, its prominence declines when the content drops to the bottom of feeds, discussion fora, and websites. Also, typically enough, art. 17 GDPR may be able to do something about viral cases within a particular social media platform, because all the content there is partially in the hands of the platform controller. With one controller, the subject can reach the complete viral chain within the platform. However, this application of art. 17 GDPR will likely get a foothold only after the outbreak, and would thus be more of a speeding up of the decay phase (that is, if it does not trigger a revival through the Streisand effect).

Where art. 17 GDPR seems to have its strongest problem-solving capacity with regard to online available information, is with regard to issues that result from a particular persistent presence of a relatively stationary reference with a low-quantitative presence in the public sphere. As such, the main functionality of art. 17 GDPR with regard to the problems identified in this study lies in addressing the establishment of a certain unwanted materialised plot by the longterm presentation of a particular reference as characteristic for the referent in one or a few signifying objects. The problematic plot created by the presence of this reference can be

reconfigured by erasing a particular signifying object, obscuring its identifiable aspects, or by reducing its accessibility and/or visibility. With this functionality, the problem-solving capacity of art. 17 GDPR is therefore best suited to deal with problems that arise on basic websites or in search engines.

Moreover, in some cases the subject can have a stronger claim to erasure. This is in cases where the content reveals sensitive information and the data subject did not make this information manifestly public herself. Additionally, in cases where the content refers to the data subject as a child, her interests in erasure are likely to carry more weight.

Lastly, it is important to remark that, as pointed out throughout this chapter, invoking art. 17 GDPR is not a guarantee for a successful reconfiguration of a subject's narrative identity. The application of art. 17 GDPR changes the informational landscape, which in turn can lead to unforeseen shifts in the narrative and/or could evoke reactions, court cases and research, thereby potentially leaving a trail of new signifying objects in its wake.

To serve humankind Art. 17 GDPR is not the only legal instrument that can be used to try to have content removed; tort law or even criminal law can be used to resolve particular cases. However, especially in its form of a counter technique to the impact of technology, it makes sense for art. 17 GDPR to work in addition or parallel to other legal measures, by first of all providing possibly a quicker solution to get content removed (despite the fact that I argued that art. 17 GDPR can be somewhat sluggish, it is still likely to be relatively fast because for instance a slander case can take years in court). Secondly, art. 17 GDPR can have an additional problem-solving effect as it triggers a duty of effort to inform derivative controllers under art. 17(2) GDPR. The strength and additional value of art. 17 GDPR therefore lies in its ability to address certain effects of the online mediation of personal information.

Adding the counter technique perspective to art. 17 GDPR has the potential to balance the interests in a manner that does justice to the interests involved, as well as to the impact of the mediating technology. It allows a fine-tuned use of the right to address particular problems, while opening up diverse possibilities to safeguard the interests of others. The key to balancing the interests in a proper manner, lies on the one hand in the gerund 'exercising' in exception (a) that concerns the freedom of expression and information (see section 8.2.9). Respecting the active societal debate, this phrasing places the emphasis on the problems that result from a persistent passive availability of content. On the other hand, it is important to also look at the intention of the 'who' that is exercising: which roles do respectively the human and the technological factor play in the hybrid intentionality that gives form to the content and the manner in which it is made present? As I have argued, the stronger the role of the technological intentionality, the stronger the case a subject should have for a successful erasure request.

I take the assessment of the respective weight of the human and the technological intentionality to be of particular importance when evaluating an erasure request, because, as I have shown, technology can easily shape someone's materi-

alised narrative identity in ways that neither the data subject nor the controller may intend, oversee, or control. Art. 17 GDPR should therefore be aimed at reducing the presence of a reference according to its accuracy and proportionality viewed in relation to the manner in which the narrative is affected on level of the narrator, the plot and/or the composition of audiences in a manner unwanted by the data subject. With the focus on technological mediation, ideally the right to erasure should have the most weight in cases where the technological intentionality shapes the narratives beyond human storytelling or expectations. This would especially be the case when the mediating technology increasingly takes on the role of external narrator, and suggests to offer a 'narrative' of people's lives. This approach does not only do justice to the rights of the individual, but also the rights of the expresser and the general public. The underlying assumption of this approach is that *human* autonomy is a fundamental value, and that, in its footsteps, we should value human intentionality more than technological intentionality — it is about what we want technology to do, not about merely going along with what it does to us. With this, I tie in to the rationale underlying the GDPR, "[t]he processing of personal data should be designed to serve mankind" (recital 4 GDPR).

Chapter 10

Conclusion

Contents

10.1 Introduction

So long and thanks for all the fish!

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, 1996

We started this journey with an answer that lacked a clear question. In the Introduction I argued that art. 17 GDPR somewhat resembles 'forty-two', the answer given by supercomputer Deep Thought to the ultimate question of life, the universe, and everything, in Douglas Adams' famous science fiction novel. The problem with the answer, as Deep Though states it, "is that you've never actually known what the question is" (Adams, 1996, p. 121). This, I argued, is the problem of art. 17 GDPR as well. It was introduced as a solution to problems that come with our current age in which online information technology is rapidly being implemented in every fibre of society. While the right appeals to everyone's imagination, it was not clear which problems the right should and could actually resolve, and how it should be applied. Rather, especially heightened by the framing as 'right to be forgotten', many analyses of art. 17 GDPR got caught up in the tension between the individual's right to privacy and the public's right to freedom of expression and information. While the discussion on the balance between values is indispensable, the practical working field of art. 17 GDPR and the role of technology in all this — the concrete issues that the right is actually supposed to address — often received little attention in the debate. With this study I therefore wanted to add a complementary perspective to the existing literature by clarifying the issues that the right can resolve. This is not to say that the balancing of rights does not need to be addressed. On the contrary. However, in order to be able to properly address the balance of interests when deciding upon erasure requests, I argue that it is vital that we have a clear view of what we are actually balancing, and what is at stake. It is important to fill these blanks and take into account the manner in which the problems are brought about and the role that technology plays therein. If we do so, the opposition that dominates the debate may be softened, or even overcome. I therefore explored the problems that the Web raises for us by assimilating our personal information and presenting it to users. In this research I set out to answer the question: *To what extent is art. 17 GDPR a viable means to address problems for individuals raised by the presentation of online personal information to Web users?*

In this final chapter, I will start by giving a summary of my main findings. Next, I conclude to what extent art. 17 GDPR is a viable means to address problems for individuals raised by the Web's presentation of personal information to users. This will be followed by a short discussion of several alternative means to address these problems. Following this, I will briefly reflect on the limits of the scope of this study. Finally, I will offer some final thoughts for the road ahead.

10.2 Summary of main research findings

The goal of the present study is to provide a better understanding of the kind of issues that art. 17 GDPR can address and how it can best be applied. In order to analyse what and how, and even if, art. 17 GDPR is a viable means to address the issues at hand, I first delved into the potential problems.

I started the exploration of the problems with constructing an analytical framework in chapters 2 and 3, that gave general insight into the relation between users, technology, information, and the referent. This framework is expanded in chapter 9, where I worked the elements into an overarching perspective. One half of the framework sees to the relation between personal information and the person to whom it refers. I explained how signifying objects can contain references to individuals. These references represent a particular piece of information pointing towards a certain referent. All the references referring to a particular person, irrespective of whether these references are correct or incorrect, form her informational persona. However, people will never see the complete informational persona of a particular referent, but only the part that is reflected by the signifying objects that they come across. Moreover, things like the shape, the quantity, and the context of these signifying objects all matter for the view on the informational persona that they afford. The signifying objects establish a certain presence of the references they contain for the perception of the user. By imbuing certain references with a stronger or weaker presence (or no presence at all), they provide people with a particular view on the referent's informational persona. This view on the informational persona plays an important role in the manner in which people understand and interact with each other: based on the information afforded by the signifying objects combined with their own background knowledge, people form a view of the referent's identity and have certain expectations of her. It is therefore in the interest of the referent to make sure that people have access to the appropriate part of the informational persona from the right perspective so that she is treated in line with her own self-image.

In chapter 9, I delved deeper into the relation between information and identity with the help of Ricoeur's theory of narrative identity. This theory concerns personal identity and leans on two pillars: on the one hand it leans on *idem* which means sameness, and on the other hand on *ipse*, which refers to the self or selfhood. Our *idem*-identity consists of a set of significations that establish a particular identity. It implies a sameness of our identity: we can be recognised as 'the same' over time, or as others. It is a 'what I am'. Contrary to *idem*, our *ipse* identity does not imply some unchanging core in our identity. The *ipse* is shaped by our reflective consciousness; we shape our personality by recognising ourselves as a certain person and by realising options of several possibilities. As such, the personal identity is never a full *idem* like an unchanging object is, but always also an *ipse*, a selfhood that is actively pursued by the individual who makes choices, expresses herself and chooses certain actions and events over others. These actualised possibilities give rise to her character. The character consists of a set of distinctive signs with which the *ipse* announces itself as *idem*, i.e. the

elements of our identity where we say 'this is who I am'. If our character traits and habits are persistent, we can see a transformation of the *ipse*, the self, into *idem*, sameness. This is where the narrative comes in. The narrative mediates the constitution of our personal identity between the character in which the *idem* and *ipse* are likely to coincide and our self-development, where we free ourselves from sameness. The narrative identity oscillates between the poles of *idem* and *ipse* and connects events by means of emplotment. Emplotment is the ascribing of a plot to a set of separate events. It is thus a 'configurational act' that mediates between the actual events and the narration of these events by organising these events in a particular manner. The character itself is a plot, a narrative construction that brings change and the permanence of our identity over time together in a whole. The narrative thus brings structure and coherence to our identity.

For the second half of the framework, I discussed how technology mediates our relation to the world, and does so in an inherently non-neutral manner. Technology allows people to perceive their world in new ways and offers them new goals that did not exist before or were impossible without technology. Connecting to the work of most notably Verbeek, I discussed that technology has a certain directionality in the manner in which it establishes a particular relation between the user and her world. This directionality is embodied in the concrete material design of the technology. While the technology is shaped by its designers, its use and effects are not limited to their intentions. Instead, the material form of the technology has an autonomous existence which itself expresses a distinctive directionality that directs the experiences and actions of users towards something. This directionality of technology is a material form of 'intentionality'. However, as technologies always play a mediating role and are dependent on their human users for the manner in which they are used and have effects, the intentionality of the technology is necessarily part of a hybrid affair of the technology and its users. As part of this hybrid affair, the intentionality of the technology does not determine how someone uses the technology, but it does co-shape the user's intention. The respective weight of the technology and the human agent in the forming of this co-shaping process, constitutes one of the crucial elements in the study presented here. Combining this perspective with the view presented on personal information, we can see that in the technology's mediation, personal information is often externalised in a particular material form. When we transfer personal information to a mediating technology, this information we materialise gains a certain autonomous existence, separate from its author. With the materialisation of personal information in the outside world, this information can tell a particular story to audiences about who we are, thereby constructing an exteriorised narrative identity of us. While technology does not construct a narrative in the classical sense, it does affect the story that is told. By presenting information in a particular manner and surrounding it with other signifying objects and processes, technology 'tells' us more than solely the content of the object. Due to this adding of extra elements and the technology's directionality, the technology takes part in the construction of a materialised version of the narrative identity. I therefore argued that in the technological mediation of the narrative, the mediating material

transfers some of its affordances and characteristics to the narrative identity that it carries. By doing so, the technology affects the role of the narrator, the audience composition, as well as the content. The mediating technology impresses some of its directionality, and thereby its intentionality, on the construction of the narrative identity itself, thereby instilling it with its own inclination towards sameness and change, affecting its meaning, and giving shape to the manner in which audiences can engage with it. The materialised narrative identity entails a complex hybrid intentionality in which the impact of the human and the mediating technology can be intertwined in different manners, with various degrees of human and technological intentionality.

With this in the background, I turned my attention to three main online applications, regular web pages, social media, and search engines, and one phenomenon, virality, while using the framework chapters as an analytical toolkit. When exploring these cases, I showed that on the Web, our social interactions and personal representations are often caught up in a battle for attention between different parties, while the mediating technologies imprint their characteristics on our online narrative identities. In this environment, human beings are constituted as a subject to particular plotlines that tell a certain story about their past and present identity. Problems with the materialised narrative identity occur on the level of the emplotment, the context of the narration, the narrator, and the audience selection. The result is that when users encode personal information online, this information takes shape in a sphere of hybrid intentionality that is co-shaped by human agents and the mediating technology. For human agents, the Web is a challenging sphere to realise the transmission of personal information in accordance with their wishes: as the Web's technology is the medium and constitutes the environment in which user encodes information, it impresses its own affordances on the processes of encoding, storage and retrieval. Everyone can publish personal information online, while the information is easily copied, edited, and transmitted. Even more, parts of the online environment are an accelerated and opaque playing field where informational distances are renegotiated in schemes of interest, popularity and profit — which can make it difficult for users to share information only in intended manners. Internet giants like Google and Facebook control important parts of the mediating technology that constitutes the online environment. As these parties have a significant interest in generating user attention for profit, they design the online environment they control in ways that play into their interests. In order to motivate users to spend attention, engage with the technology and provide content, they manage information flows in which meaning is often tied to popularity and advertisement revenue, while they keep the underlying mechanisms of the technology hidden from users.

With the Web's publishing and transmission affordances and the control over the mediating technologies in the hands of various players, the referent's control over the *who* with whom information is shared, the *when*, the *how*, and *the extent* to which this information is shared, can all be challenged. The signifying objects shaped in this realm of various constructions of hybrid intentionality can easily be constructed and reassembled in such a manner that the presented persona

becomes an exaggerated, distorted and problematic reflection of the current selves of the referents. Online technologies like search engines can actively affect the emplotment of the materialised narrative. The broad informational scope of the Web combined with centralised information flows can easily lead to the construction of a relatively 'flat character'; differentiations between time, space, audiences, context, relations, and public and private are diminished and replaced by differentiations based on association, interest and/or popularity. Meanwhile, the impact of the online persona on our offline lives is intensified as our current Zeitgeist is characterised by being an information society which heavily uses ICT for virtually all aspects of life. The consequence is that our (potentially problematic) online persona plays an important representative role in many parts of our lives: the high frequency with which we engage with the Web leads increasingly to a setting where we *are* our online representation in interactions.

The online mediation of personal information can trigger several issues for the subjects of this information. I will now briefly list the main issues per case-type.

Web pages In its mediation, the Web leaves an imprint on the online narrative identity: it imbues people's narrative identity with a scale, scope, durability, flexibility and access speed that can easily result in a longterm and global presence. The potential issues already start with the encoding as the Web gives rise to interactive narratives to which anyone can contribute, human as well as technological others. Even content that is processed with the subject's consent can lead to problems, because the subject may overlook or misjudge the implications of a particular instance of online processing. Once online, content is embedded between the Web's other content and becomes open for hyperlinking and further processing, and takes on the Web's access characteristics. The content can be accessed by an audience from anywhere at any time and can be shaped in myriad ways along the trail of hyperlinks. The online narrative is therefore dynamic and subject to new cycles of retrieval, disclosure, dissemination, combination and collisions with other online narratives. Meanwhile, audiences are difficult to escape and audience segregation failures are easily made, especially given the fact that plain web pages tend to be publicly accessible by default. Several problems can therefore occur for the online narrative identity: the self-expressions can inadvertently become material representations of the individual's character due to the access and storage afforded by the Web, while as a result of the combination of the dynamic affordances and the highly communicative and networked character of the Web, the materialised narrative identity is easily hijacked by human and technological others who add and remix content and/or present it to unintended or unforeseen audiences. The online externalised narratives can be problematic because they may attribute a certain quality to the self that the subject may not consider to be representative for her, or at least less than this narrative suggests. As others respond to the subject based on this information, they can open or close certain options to the subject, like offering her a job or not, starting a friendship or keeping more distance from her. Meanwhile, these reactions of others can also reflexively affect the subject's self-perception. Additionally, being confronted

with the content herself can affect the subject's self-perception by, for example, triggering dormant memories or providing her with an unexpected view of herself.

Social media As a new realm for social interaction, social media have led to a huge wave of materialisations of personal information. Due to its social and self-broadcasting character, they are regularly used by users to present themselves, experiment with their identity and associate themselves with particular others, ideas, or subcultures. Combined with their easy editing options and a strong focus on the now, social media therefore have their weight in the construction of the self by means of self-presentation. However, due to the participatory role of the audience, the audience can highly impact the narrative that is told by adding and annotating content and co-shaping its meaning. This results in a participatory personal narrative that is shaped in the social medium's triad intentionality of the subject herself, others and the mediating technology. Meanwhile, the medium's often opaque architecture entails a constant risk for audience-segregation failures. The core of the issues on social media consists of an overrepresention or inadequate image towards different and potentially unintended audiences in the here and now — a deformed representation of the self shaped in the social medium's triad intentionality.

Search engines Search engines appropriate content of others and give their own spin to it: by selecting, framing and organising excerpts of original objects in a particular manner, the search engine emplots a set of references originating from multiple narrators into a new narrative and sets the context and audience for the story. By doing this, they take up the role of a new narrator and tell an audience what is valuable. Meanwhile, their pivotal role as gatekeeper on the Web establishes them as narrators with a rather authoritative voice. As the search engine construes a narrative based on request, it turns the search string into the topic of the search results and may bypass the intentions of the original authors of the publications, potentially decontextualising content and revealing content to unintended audiences. The search engine's algorithmically driven plot can easily portray a particular reference as part of an individual's character by always presenting it as a top result. Moreover, the plot can misrepresent the individual if it displays erroneous or decontextualised references, or certain references can be given a salient role in the personal narrative, while they, in fact, do not revolve around the individual, or only do so in a marginal manner. Additionally, search engines may nudge audiences towards particular narrative angles by means of autocomplete. Combined with our dominant use of their services, the narratives constructed by search engines are likely to play a defining role in the construction of the online narrative identity.

Virality A viral event has a strong impact on an online personal narrative: a single reference becomes so strongly present that it turns into the main plotline of a subject's complete narrative. The other signifying objects are likely to be sucked up

within this plot created by the viral reference. The viral plot may even establish a leading narrative that caricatures the referent to the extent that she may become a longterm symbolic character. Reconfiguring a narrative shaped by a viral outbreak requires a fundamental change in the overall emplotment. Viral cases potentially also cover all the issues connected to the various applications discussed above, because the content is likely to be spread across diverse applications.

The conclusion of the problem analyses, is that there is not one problem. Instead, there is a group of problems that have a certain family resemblance in the fact that they influence the online materialised narrative identity beyond the wishes and/or expectations of the subject. The 'long-lasting memory' of the Web, the kickstarter of the wish for a 'right to be forgotten', is just one of the issues. A more important generic trait in this family of problems is the Web's connectivity in combination with the affordances of digital objects: content is easily created, edited, and spread beyond expectations. Online personal information can lead to diverse representative problems for the subject, and vary in the factors that play a main role in the manner in which they come to be. The problems revolve to a great degree around the proportionality of a particular reference in relation to the plot of the narrative identity as the subject believes she should be represented. For online content to cause complications, it does not even need to have an extraordinary or negative character. As we have seen in the BBC cases, not all cases involve extreme events, and many of them are even quite mundane and fall into a broad range of relatively common topics.

Overall, the mediating technology plays a prominent role in many of the identified issues. However, it is not just the technological mediation that plays a major role in this: the human agent is often an accomplice to the problems. While humans who process information online are always expressing a hybrid intentionality, their decisions play an important role in this processing. Online, people are encoding new content, as well as copying, editing, and remixing existing content. Such creation of new signifying objects and descendant objects is often simplified by online service industries that offer simplified options for the encoding of information with the use of WYSIWYG website generators like Wordpress, and social media like Instagram, Facebook, and Twitter. In these cases, the encoding and dissemination of content is industrialised to a greater or lesser degree. Despite the role of the mediating machinery and its often limited transparency and control options, users are the ones that make the final choice in the encoding and dissemination of this content. Human encoded content that itself is experienced as problematic by the referent (these range from singular encodings to a viral outburst), is therefore a problem that for an important part can be attributed to human intentionality.

Capability of art. 17 GDPR to address the issues The next phase in this study was to assess whether art. 17 GDPR is capable of addressing the identified issues. However, this was not a simple case of applying the article. In chapter 8, I discussed that the right was caught up with some challenges of its own. The

main problem of the right is that it seems to suffer from a lack of clarity that could at least partially be attributed to the right's double naming and framing. Moreover, while 'forgetting' in relation to art. 17 GDPR can serve as a useful concept, the conceptualisation of art. 17 GDPR as a 'right to be forgotten' overall seems to steer our view in suboptimal directions. It clouds the impact of the Web on other factors that affect the presence of personal information, such as space and proportionality, while its conceptualisation is highly metaphorical. I therefore argued that it is better to take the mechanisms of the right as a point of departure for our further investigation. The mechanisms of the article express a particular functionality: they are focused on giving the individual a certain degree of control over her personal information by means of erasure. However, even 'erasure' in the context of art. 17 GDPR seems to have a somewhat metaphorical character: 'erasure' can take on diverse forms, some of which are technically not even forms of erasure, but of blocking. The forecasting case law of the *Google Spain* case paved the way for this, as well as for the right's ability to address descendant objects while leaving the original in place. In spite of the possible vagueness that may come with such a metaphorical understanding of erasure, it is exactly by allowing this broad scale of various forms of erasure, that the right can offer a lot of room to manoeuvre in order to resolve the problems while doing justice to the various interests involved. Erasure could entail a full or partial erasure of the content itself, an erasure of the content from view for (particular) users, or by implementing certain manners of processing that reduce the presence of a particular reference.

In chapter 9, I delved further into the fact that art. 17 GDPR is not an absolute right that gives the individual full control over all the processed information relating to her. Instead, the interests of the subject need to be carefully balanced against the interest of the controller, a potential original publisher, and the general public in the ongoing processing of the information. I argued that in order to come to the most balanced solution, we should fine-tune the form of erasure specifically *to the impact of a particular technological mediation* on the produced narrative. Advancing this perspective further, I argued that art. 17 GDPR should help data subjects to reconfigure the technologically mediated narrative so that the subject's freedom to construct her own narrative identity is protected against unreasonable constraints raised by the processing of personal information. Art. 17 GDPR should therefore be aimed at reducing the presence of a reference according to its accuracy and proportionality viewed in relation to the manner in which the narrative is affected on the level of the narrator, the plot and/or the composition of audiences in a manner unwanted by the data subject. As such, it can help individuals to reconfigure their materialised narratives when these are disproportionally shaped by a technological intentionality or at the hands of others who have no preponderant legitimate interest to tell these stories about us. Given the important role of the mediating technology in this, art. 17 GDPR should ideally have a strong focus on the impact of the technological mediation on the narrative identity. Art. 17 GDPR can be highly valuable if we understand it as being able to function as *counter* technique to counteract the impact of technology.

Understanding art. 17 GDPR as being able to function as a 'counter technique' is a fruitful complementary understanding of the right: it has the potential to balance the interests in a manner that does justice to the interests involved, as well as to the impact of the mediating technology. It allows a fine-tuned use of the right to address particular problems, while opening up diverse possibilities to safeguard the interests of others. The open interpretation of erasure allows us to look for an application of erasure that is proportional to the interests of others, and that also alleviates the problems for the subject. With the focus on technological mediation, ideally the right to erasure should have the upper hand in cases where the technological intentionality shapes the narratives beyond human storytelling, intentions and expectations. This would especially be the case when the *mediating technology* increasingly takes on the role of external narrator, and suggests to offer a 'narrative' of a subject's life. By taking human intentions as an important guiding principle in assessing whether a certain technologically mediated presentation of online information should be addressed, this approach places the focus on the value of *human* autonomy. It also ties in with two important points of the GDPR: its rationale and its focus on the purposes of the controller. By placing human intentions centre stage, understanding art. 17 GDPR as capable of addressing an overly strong impact of technological intentionality supports the GDPR's rationale that the "processing of personal information should be designed to serve mankind" (recital 4 GDPR). Additionally, by taking human interests as a guiding principle, this approach connects to the GDPR's overall focus on the goals of the controller: with the legality of the information processing depending on the purposes of the controller, her intentions with regard to the processing of the information play an important role in the balance of interests.

The subject-induced character of the right connects to the value of human autonomy: art. 17 GDPR imbues the referent with a significant autonomy with regard to exercising control over her online narrative identity by allowing her to choose which controller to target with an erasure request and when to request the erasure. She can thus exercise a certain control over her materialised narrative identity. With this she can aim to address issues that result from the unwanted processing of her information by others, or by the unforeseen impact of technological mediation that followed, for her unexpectedly, from the processing of her information with her consent. However, it is this same subject-induced character that places certain limitations on the right's effectiveness: the subject may need to invoke art. 17 GDPR for every case in which a particular reference is mentioned online — and she may need to do this quickly. It is questionable whether the referent in all cases will be aware in time of the content's existence to prevent the worst of the problems, or is even able to locate all the occurrences of the content. The subject-dependent invoking mechanisms therefore seem somewhat sluggish to address the highly dynamic Web environment where everyone can copy, upload and disseminate content worldwide in seconds. Invoking the right to erasure in the case of a reference with a high quantitative presence will be burdensome for a single individual, and addressing a full viral outbreak will even be impossible for an individual to achieve. The inability of the right to address

a viral outbreak is unfortunate, because the impact on the subject's life is high. However, art. 17 GDPR may be able to do something about viral cases contained within a particular social media platform, because all the content there is partially in the hands of the platform controller. With one (joint) controller, the subject could reach the full viral chain within the platform.

Despite the problem-solving potential of art. 17 GDPR, it is not suitable to address in a satisfying manner all the problems that I touched upon. Next to its limited capability to address a viral outbreak, art. 17 GDPR is also unsuitable for cases where the individual herself errs as a controller because the right cannot be invoked with regard to content that is under her own control. Additionally, the right has difficulty to address cases in which the individual does not want to sever the ties between her and the object, but instead is interested in specifically setting certain audiences for the content or preventing misinterpretations. In these cases, the right to erasure can be of limited help to the subject. One of the pivotal issues that the right cannot satisfyingly address with its erasure functionality, are audience segregation failures. Also, the functionality of art. 17 GDPR for social media can be hampered by the household exemption. While the household exemption in theory should work well together with art. 17 GDPR because the right ideally addresses the (semi)publicly available or organisation controlled narrative identity of individuals, it does potentially place a significant use of online technology outside of the right's functional scope. The risk of the household exemption, given the connective character of the Web, and especially the high connectivity of social media users, lies in the fact that the 'household' character of online information can be rather volatile. A signifying object that is initially only visible to twenty people, can with a few clicks become visible to thousands of users. However, the household exemption is an important exemption to protect individual autonomy with regard to the use of a more personalised externalised memory: individuals may process all sorts of personal information about others for personally relevant reasons. The risk that some of this information may spill outside the household scope is therefore one that is well worth the interests that it protects, but this does likely leave some problems unattended to — at least by art. 17 GDPR.

Moreover, even when successfully invoked, erasure as result of an art. 17 GDPR request is not a guarantee for a favourable reconfiguration of a subject's narrative identity. The application of art. 17 GDPR changes the informational landscape, which in turn can lead to unforeseen shifts in the narrative and/or could evoke reactions, court cases and research, thereby potentially leaving a trail of new signifying objects in its wake.

Based on all of the above, I concluded in chapter 9 that art. 17 GDPR is best equipped for addressing issues that result from a particular persistent presence of a relatively stationary reference with a low quantitative presence in the public sphere. As such, the main functionality of art. 17 GDPR lies in addressing problems that result in the establishment of a certain unwanted narrative plot by the presentation of a particular reference in one or a few signifying objects. The problematic plot created by the presence of this reference can be reconfigured by erasing the content,

obscuring its identifiable aspects, or by reducing the accessibility and/or visibility of the signifying object. With this main functionality, the problem-solving capacity of art. 17 GDPR is best suited to deal with problems that arise on basic websites or in search engines. Of these, the ability of art. 17 GDPR to address content in search engines is its biggest asset. Of all the online applications examined, search engines have the biggest impact on our narrative identity and generally the strongest expression of technological intentionality therein. Meanwhile, search engines are also the most successful point to apply art. 17 GDPR in order to address problematic narratives spread over the Web. In guiding users towards content, search engines can increase or decrease the chance that a specific user comes in contact with a specific reference. They thus have a powerful influence on the formation of audiences for particular content and can reduce the presence of certain online references. The strong position and gatekeeping role of search engines therefore imbues them with the power to address a fair share of the issues, or at least address the issue in its most intense form. The erasure of search results can even help to mitigate some of the problematic impact of a viral information flow on its subject.

10.3 'Forty-two' revisited: conclusions

With the study presented here, I aimed to fill what I perceived to be a gap in the debate surrounding art. 17 GDPR: a clear view on the problems that it can address. The study provides us with a sharpened view on when and how to apply art. 17 GDPR. By providing an in-depth view of the how and what of the problems, as well as the respective roles of technology and human agents herein, the analyses can help to contextualise, and in some cases maybe even overcome, potential conflicts between the right to erasure and the other interests involved, most notably those protected under the right to freedom of expression and information. Moreover, by clarifying how the problems come to be, and what the key elements in the various problems are, the analyses provide handholds for a precise and fine-tuned application of erasure in a manner that respects the various interests at stake. With this in the background, we can now answer the main question: *To what extent is art. 17 GDPR a viable means to address problems for individuals raised by the presentation of online personal information to Web users?* In this section, I will draw conclusions about the extent to which art. 17 GDPR can address the identified problems for individuals in a viable manner.

To start with, we would do best to move away from approaching art. 17 GDPR as a 'right to be forgotten'. When I argued in one of my first papers that it would be beneficial for art. 17 GDPR to "draw more heavily on the mechanisms of human forgetting" (Korenhof, 2013, p. 126), I fell into the same 'name' trap as many others when it comes to understanding art. 17 GDPR. This is not surprising, because a 'right to be forgotten' appeals to the imagination. Also, the potentially persistent and highly detailed memory capacities of the digital milieu are indeed one of the problem causes that we need to deal with in contemporary life. However,

as I have shown, this perspective often does not fully nor accurately reflect the issue at hand. Approaching art. 17 GDPR from the 'right to be forgotten' angle can obscure a part of the problems and puts us at risk of misidentifying them, while at the same time complicating and muddling the discussion about art. 17 GDPR. To repeat a catchy phrasing from the discourse surrounding art. 17 GDPR: we should forget about the right to be forgotten (see e.g., Cunningham, 2017; Lynskey, 2013). However, I argue that we should only forget its unfortunate name, and not the workings of the article.

As set out in this study, in the online mediation of personal information, the individual is produced as the subject of a materialised narrative. This narrative may attribute qualities to the self that the individual may not consider to be representative for her, or less so. Moreover, due to the highly present role of the Web in contemporary society, the individual's identity can easily be dominated by online materialised narratives, potentially leaving her little room to construct her own self-identity as she sees fit. The potential fear of content becoming long-term publicly accessible may even give rise to self-censorship. By allowing data subjects to request the erasure of particular content under certain circumstances, art. 17 is an instrument that empowers individuals by giving them a certain degree of control over their materialised narrative identity in an era where that narrative is evermore materialised and shaped by others — both humans and technologies — rather than by the self. As art. 17 GDPR can function as a means to counter certain instances of information processing that shapes the materialised narrative identity in a manner unwanted by the data subject, it can specifically be employed as a *counter technique*. As a counter technique, the right should aim to counter the unwanted impact of a mediating technology on the materialised narrative by means of erasure. Additionally, paragraph (2) of art. 17 GDPR, which aims to account for the implications of the multiplication and transmission affordances of online digital information by requiring the controller to inform third parties that the data subject requested the erasure of the content, makes sense when we approach the right in the perspective of aiming to function as a counter technique in the internet era.

The goal of the right is to help protect a certain degree of autonomy of the subject over her informational persona. In this context, the subject-induced character of the right works well with its goal and many of the issues that can ideally be addressed with the right: it is the subject herself that is given a certain degree of control to try to adjust her materialised narrative identity. This is especially valuable in the light of the role of the internet giants like Google and Facebook, who as information controllers easily dominate how referents are produced and presented as a subject to others in the online milieu. The right thus somewhat compensates for this power imbalance by allowing individuals to get a certain degree of control over their identities in face of the internet giants. However, the right also respects the interests of others in the information processing. Whether or not content should be removed in response to an art. 17 GDPR request highly depends on the interests that the controller, the general public, and the potential original publisher have in the particular processing of the

information. However, as I argued in chapter 9, we should assess this balance with the focal point on the interests as they relate to human intentionality. Because the right allows various forms of erasure, we can look for nuanced ways to provide the data subject with a certain autonomy over her self-presentation, while also respecting the interests of others.

My main conclusion is that, ideally, the right to erasure should have the most weight in cases where the technological intentionality affects the narrative identity beyond the expectations and intentions of the human narrator(s). In this capacity, art. 17 GDPR mitigates the problematic imprints of the technological intentionality on our narrative identity, while respecting relevant human intentions with an eye on its value for the controller, a potential original publisher, and the general public. The wide interpretation of erasure afforded by art. 17 GDPR can be used to do justice to the varying value of information in different contexts, and to the impact of the technological mediation. By focusing more on tinkering with the processing of content, and less on the full erasure of objects, information will in many cases remain accessible, but in a better contextualised manner. This will be beneficial for a proportional symbolisation of the individual by her online personal information, while also taking into account the public interests and the (human) intentions of all parties involved.

Overall, art. 17 GDPR is a viable means to address a fair share of problems for individuals raised by the presentation of online personal information to Web users. Its strong suit lies in addressing issues that result from a particular persistent presence of a relatively stationary reference with a low-quantitative presence in the public sphere. As such, the main functionality of art. 17 GDPR lies in addressing problems that result from the longterm presentation of a particular reference as characteristic for the referent in one or a few signifying objects. The problematic plot created by the presence of this reference can be reconfigured by erasing a particular signifying object, obscuring its identifiable aspects, or by reducing its accessibility and/or visibility. With this functionality, the problem-solving capacity of art. 17 GDPR is therefore best suited to deal with problems that arise on basic websites or in search engines.

However, art. 17 GDPR is unable to address all issues in a viable manner. While being beneficial for the subject's autonomy, the right's subject-dependency comes with a disadvantage: while in theory the right is a fit instrument to address a significant number of issues, its practical usability is restricted by the (especially digital and legal) skills and resources of the subject. More skilled people will be more proficient in finding and contacting controllers and make a stronger case for erasure, or in arranging for an NGO or the like to help them invoke their right. This may increase the digital divide with regard to control over online narratives. Another consequence of art. 17 GDPR's subject-dependency is that the right is difficult to use for the erasure of a high number of signifying objects in the hands of different controllers. This renders art. 17 GDPR rather helpless to address a narrative identity that suffers from a broadly published or even viral reference. This is unfortunate, because virality raises the kind of problems that would ideally be addressed due to their major impact on the subject's narrative identity.

These practical restrictions to the usability of the right reflect one of its weak points: while art. 17 GDPR is an able instrument to address a part of the issues resulting from the online presence of personal information, it lacks the heavy technological backbone of the technologies that realise the information processing that it is supposed to counteract. As a legal instrument, art. 17 GDPR is a counter technique, but not a counter *technology* — a practical technological tool. It is thus a different type of instrument than the technological information processing that it aims to address. Although the right to erasure helps to create a better power balance between data subjects vis-à-vis the information controllers on the legal level, it does this in a milieu in which these controllers control and generally better oversee the mediating technology. A technical power imbalance between data subject and controllers thus remains. This is tricky because the mediating technology plays a key role in many of the issues that are raised by the presentation of online personal information to Web users. Novotny and Spiekerman therefore already pointed out that this is a potential problem for data subjects (Novotny & Spiekermann, 2014). In section 10.4, I will discuss several instruments that have been suggested by scholars to be used next to, or instead of, the right to erasure, inter alia in order to deal with this technological power imbalance.

Moreover, art. 17 GDPR's functionality in the form of erasure, even with its wide scope of possible applications, is not a suitable means to address all issues. A prominent example of this, is the right's incapability to address in a satisfactory manner current audience segregation issues caused by the data subject herself or by others. Erase will in many cases defy the purpose of the publication altogether. Despite the fact that by sharing certain information in a particular context of human action plays a pivotal role in the creation of audience segregation issues, sharing the information with the wrong audience is not the *intention* of the human author. While in some cases human stupidity may be to blame, in other cases the technological mediation plays a significant role by obscuring the publication conditions to such a degree that the human author can hardly be blamed for misinterpreting the scope of the audience. However, this does not mean that art. 17 GDPR cannot address any audience segregation issues at all. The right can battle potential audience segregation failures when these are the result of descendant objects being published in a secondary location: by erasing these objects, the issue would be resolved. Unfortunately, given the high speed of information distribution and access in the online environment and the retroactive character of the right, the segregation failure likely already took place before the right was invoked — and the descendant publication may already have caused unfortunate effects for the data subject.

Art. 17 GDPR is not the answer to life, the universe, and everything. It also is not the answer to all the problems raised by the presentation of online personal information to Web users: there are limits to what art. 17 GDPR can do. Despite these limitations, art. 17 GDPR is overall a viable means to address a fair share of issues. Especially its capability to address content presented in search engines, which can help to mitigate the impact of problematic narratives on the Web, at least in their worst form. However, it is important to point out that while the

right can address several issues, it does so by mitigating mainly the symptoms of the problems. With its focus on *existing* content, art. 17 GDPR addresses that which is already processed by the technology. As such, the right to erasure changes the outcome of a certain type of processing, and not the steps that led thereto. For instance, when a particular search result is removed in response to a right to erasure request, this does not change the algorithms producing the results and the priority given to particular sources. Only a specific search result is removed — or rather blocked — in response to a particular query. I therefore argue that the right somewhat resembles a painkiller, instead of a cure to a disease or wound. Like painkillers, the right may be able to battle the symptoms long enough for the wound to heal, or the disease to pass over. Unfortunately, not in all cases do wounds heal or diseases just go into remission by themselves. Moreover, diseases may return. The same counts for objects that were erased. Due to the replicative and networked affordances of the Web, art. 17 GDPR's curing effect may be short-lived as content can be uploaded again, or search results can by reintroduced by changing the URL of the source. In this sense, the Web is comparable to a living body, which grows, replicates cells, declines, and changes over time. With its *ex post* functionality, the right to erasure addresses a pain that already exists. Since the pain may get much worse without treatment, the right to erasure can thus be a viable means to address some of the pains that come with the online presentation of personal information to users. However, painkillers come with risks: they harm the body, cause addiction, or, in the worst case scenario even wreck it to death. The right to erasure also comes with risks: it can be misused and it may reorganise the informational landscape in possible unforeseen and unwanted manners. Yet, these risks should be relatively easily mitigated because a proper balance of interests should prevent misuse, and some prospective planning should prevent data subjects from being confronted with new problems caused by a reorganised informational landscape.

The right to erasure can thus help to mitigate the worst of the pains. Painkillers are needed because doctors and good lifestyles do not prevent all diseases or injuries. However, the use of painkillers may also be a signal of a more serious underlying cause, something that does not easily pass. The inventor of the Web, Berners-Lee, voiced in 2018 that there is something fundamentally wrong with the Web: the current shape of the Web embodies an unbalanced power structure in which users lack control over their own information. He therefore states: "Today, I believe we've reached a critical tipping point, and that powerful change for the better is possible — and necessary"[1]. Whether the GDPR succeeds in addressing this unbalanced power structure with regard to personal information in a viable manner remains to be seen. If we see an ongoing prevalence of erasure requests over time, this may indicate that the preventative medicine of the GDPR in total does not work quite as it should, or that the GDPR is not enough to realise a 'healthy Web'. In this case, we may need to consider looking elsewhere for a cure. Time will tell.

[1]Tim Berners-Lee, "One Small Step for the Web...", *Medium*, 2018. `https://medium.com/@timberners_lee/one-small-step-for-the-web-87f92217d085`, last accessed 19-03-2019.

10.4 Other means to reconfigure the online narrative

The use of art. 17 GDPR is not the only possible way in which we can (try to) address the issues that I identified. For other ways to deal with the issues, we can look in the direction of three of the main forces that have been discussed: law, human beings, and technology. In this section, I will give some examples of alternative ways to deal with unwanted narrative identities on the Web.

Starting with law. Law provides several alternatives to art. 17 GDPR that may be viable means to address certain issues with the online narrative identity. Examples of such alternatives can be found in the GDPR itself. I will discuss three alternatives to art. 17 GDPR that legal scholars have found in the GDPR. The first one is art. 5(1) GDPR. Graux, Ausloos and Valcke recognise a 'passive' (at least from the perspective of the data subject) right to erasure in what is now art. 5(1)(d) GDPR, which sees to the accuracy of information (personal information shall be "accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay"), and 5.1(e) GDPR which entails a storage limitation (personal information shall be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed") Graux *et al.* (2012). The advantage of these provisions is that they bypass the problems that result from the subject-induced character of art. 17 GDPR. However, Graux, Ausloos and Valcke warn that the power of such passive safeguards should not be overestimated because the actual implementation of these safeguards will be dependent on the enforcement of the law (Graux *et al.*, 2012). For the enforcement of the GDPR, Data Protection Authorities play an important role. Data Protection Authorities are independent supervisory bodies that check whether controllers and processors comply with the GDPR. If a controller does not comply with the GDPR whilst falling within its scope, the Data Protection Authority can issue a fine. However, given the sheer scale of information processing actions and controllers that fall under the GDPR, it is unlikely that Data Protection Authorities are able to oversee and check all corresponding controllers and processors. It may therefore be wise to not overly rely on the passive counterpart of the right to erasure.

Another GDPR right that might be able to address certain cases, is the right to object, art. 21 GDPR. Ausloos argues that the right to object might be a more empowering right for data subjects than the right to erasure (Ausloos, 2018, p. 369). Art. 21 GDPR allows data subjects to target processing operations instead of particular content (Ausloos, 2018, p. 94). It provides users with a means to put a 'stop' to these processing operations, but does not necessarily require the erasure of content that has been processed thus far. The right to erasure and the right to object can therefore be used to complement each other, depending on whether the data subject wants to have particular content erased or only want particular processing operations to stop (Ausloos, 2018, p. 369). Ausloos gives the example

of a data subject who wants to retain her social media profile, but wants to put a stop to the processing of this information for advertisement purposes (Ausloos, 2018, p. 369). With the help of art. 21 GDPR she can try to put a halt to this specific processing, without having to erase her profile information with the help of the right to erasure. However, I look at this case a bit differently. I argue that the processing of profile information for advertisement purposes is generally performed with a copy of this content (especially as this processing is often performed by a third party). With art. 17 GDPR the data subject could target this particular descendant object and have it erased, while her original profile information remains equally intact. Which of the two rights is in the end preferable for this case, will depend on future interpretation of the corresponding rights by Data Protection Authorities and courts. However, in cases where no copy is made, art. 21 GDPR can be a viable means to stop certain processing actions, like for example stopping content from being edited, while leaving the original content intact. I therefore agree with Ausloos that the combination of the right to erasure and the right to object could allow data subjects to more precisely influence the manner in which controllers process their information in a way that is fine-tuned to their wishes with regard to their online narrative identity, even if how they complement each other in practice still may need to pan out a bit more. Readers interested in a detailed research into the relation between art. 17 GDPR and art. 21 GDPR, I would like to refer to the research by Ausloos (2018).

Another GDPR alternative is art. 18 GDPR, the right to restriction of processing (Ausloos, 2018, p. 369). This right allows the data subject to request the controller to restrict the processing if she contests the accuracy of the information (art. 18(1)(a) GDPR). This restriction is in force until the controller verifies the accuracy. The data subject can also request the controller to restrict the processing if the processing is unlawful, but the subject does not want the controller to erase the content (art. 18(1)(b) GDPR), or if the purposes for processing expired, but the data subject needs the information for exercising of her legal claims (art. 18(1)(c) GDPR). Last, the data subject can request to restrict the processing pending the verification of the controller's grounds in an art. 21(1) GDPR claim (art. 18(1)(d) GDPR). As Ausloos rightly points out, due to the temporary nature of art. 18 GDPR, it will be of little help to definitively address an issue. This is therefore only a suitable alternative for cases in which a temporary stop of the processing is needed and the processing can continue as-was after the temporary break. In the context of this study, I did not come across any case where this may have been a preferable means over the right to erasure to address the issue.

Besides the GDPR, there are also other laws that may be of help with addressing particular online content. Examples are copyright laws, child protection laws, and slander and defamation laws. These laws see to specific cases and are directed towards particular content that is problematic or protected. While these laws in theory do not require erasure as they merely state that someone should pay a penalty for having processed the information, in practice they will result in most controllers choosing to erase the content because they are not willing to pay

future fines. However, because the scope of this study was specifically restricted to cases that did not evidently fall under these laws, I expect these laws to be of help only in a part of the cases, but not all. These laws will leave much of the online content untouched — content that can potentially be targeted with art. 17 GDPR. Additionally, the content that is covered by for example criminal law may still be more quickly addressed with the right to erasure because these cases can take a long time in court.

The last example (bordering) on the legal front that is worthwhile to mention, is a proposal by Powles and Floridi that is roughly the opposite of art. 17 GDPR, but aims to address the same kind of problems. They propose to give people a right that involves adding more information instead of less. Powles and Floridi argue in favour of a 'right to comment': by allowing people to add comments to, for instance, search results, a more contextualised picture can be presented (Powles & Floridi, 2014). By adding 'sticky notes' and commentaries to signifying objects (e.g., saying that that the information has been proven to be incorrect, or that it is outdated), decontextualisation issues can be reduced or even resolved. This could be helpful with regard to content in online archives because it would provide users with an accurate view of the past in relation to the present. The adding of contextual information is a way to achieve a compromise between competing interests. However, it should be noted that additional information is likely unable to fully undo any first impressions made by the original (outdated or incorrect) content (de Mars & O'Callaghan, 2016, p. 280).

While art. 17 GDPR is thus certainly not the only legal provision — not even in the GDPR itself — that may address issues with the online narrative identity, I argue that despite the existence of alternative means to achieve a similar result, art. 17 GDPR does have its merits. In its form of a counter technique, it makes sense for art. 17 GDPR to work in addition or parallel to other legal measures, first of all by providing in some cases a possible quicker solution to get content removed (because, for instance, a slander case can take years in court). Secondly, art. 17 GDPR can have an additional problem-solving effect because it triggers a duty of effort to inform derivative controllers under art. 17(2) GDPR. The strength and additional value of art. 17 GDPR therefore lies in its focus to address certain effects of the online mediation of personal information. While many of the other legal measures tend to focus on the assessment of the human-intended content, art. 17 GDPR is able to more quickly curb specifically the undesirable nature of the technological mediation.

Of course, invoking law is not the only possible way to deal with problematic online narratives. This brings me to the second force that may be used to address the issues: human beings. Users may be able to resolve some issues by taking control over their narratives by themselves. A first example is given by Dholakia and Zwick. They found that some users attempt to gain some control over the externalisation of their personal information by means of 'ultraexhibitionism' (Dholakia & Zwick, 2001, p. 10). This could reduce the subjection of the subject to an identity that is impressed upon her by others, or at the very least give the subject some feeling of control. While trying to grab control over the own narrative

may certainly help, doing this by revealing ever more information may easily lead to new problems as the data subject may reveal too much, or her publications may be picked up by others and become decontextualised in some way. Where Dholakia and Zwick refer to ultraexhibitionism as an act of resistance, which may have been a plausible position in relation to the Web around 2001 when they wrote the text, anno 2019 it seems more like an act of desperation if it is done in an attempt to regain control.

A similar technique but with a different angle, is Brunton and Nissenbaum's suggestion to make use of obfuscation: the deliberate adding of more, ambiguous and confusing information, in order to interfere with the dominant narrative that the information collection tells about us (Nissenbaum & Brunton, 2015). While this may indeed drown out the presence of problematic references, the question is whether overall this leads to views on the subject's informational persona that are helpful for her narrative identity: due to the obfuscating content, she may risk to come across as vague or ambiguous.

Something in between these two subject-activated tactics is the tactic used by 'reputation managers'. Reputation managers aim to reshape an individual's online portrayal by adding an abundance of prominently present positive information to the Web, thereby drowning the visibility of the targeted information (Ronson, 2016, p. 194). While this tactic can help to change the main plotline of the online narrative, depending on the added content, it entails the risk of representing the subject as being relatively boring and/or plain, or maybe make her look like she is arrogant or likes to brag a lot. Additionally, a more important drawback, is that reputation management costs money, and will largely benefit the well-off.

What is striking, is that all these tactics that aim to interfere with the current online presented narrative identity, depend on the encoding of more information. Moreover, unsurprisingly, being rooted in the referent to take action, the success of the solution highly depends on the data subject. Like art. 17 GDPR, these tactics are therefore subject to some of the same inefficiency problems as art. 17 GDPR (unless the subject hires a reputation manager). That this is indeed a problem, was ascertained by Novotny and Spiekerman who researched the user wishes and needs with regard to personal information. In this research they found that with the increase of personal information on the Web, "[u]sers are already reaching the limits of manual control over the [personal information] they have disclosed over time" (Novotny & Spiekermann, 2014, p. 12). Ideally, users therefore should get the support of technology that helps them to locate content, identifies critical information, and suggests countermeasures (Novotny & Spiekermann, 2014, p. 12). This brings me to the third force that we can consider to use to resolve the issues: technology itself. Applying technology to counter a certain unwanted impact of technological mediation would truly serve as a *counter technology*. This would address the technological imbalance that is now one of the more critical weaknesses of the right to erasure.

An example of a counter technology, an idea proposed by inter alia Mayer-Schönberger, is to build in 'expiry dates' into digital information (2009). Several researchers, groups and organisations worked on different technological implemen-

tations of such a built-in auto-destruct. An example of this is the project 'aging file system' (Accapadi & Weber, 2011). The aging file system is a methodology that aims to realise aging in digital files similar to the manner in which paper and physical photographs age. The methodology is implemented in a file system and affects the files therein (Accapadi & Weber, 2011). The files are aged based on "parameters like ambient temperature, rate of aging, simulated type of paper or photo paper" (Accapadi & Weber, 2011). While the aging file system ties the aging of files to the properties of their physical counterparts, one can also imagine a similar aging system based on the passing of 'social time', like basing the aging rate of files on the age of the data subject.[2]

The upside of applications that would enforce automatic technologically induced erasure over time, is that all the problems and risks of the subject dependency of art. 17 GDPR are bypassed. However, automatic deletion in response to a set expiry date may come with drawbacks. On the one hand, with automatic erasure, we may risk throwing the baby out with the bathwater: too much content may easily auto-destruct. On the other hand, as Koops argues, the difficulty of making well informed decisions with regard to expiry dates and the fear of possibly employing a too-quick automatic erasure will likely result in people setting long expiry dates, so that overall the setting of expiry dates would not solve much (Koops, 2011, p. 242). It is therefore worthwhile to also consider technological solutions that connect to the upside of subject-induced erasure, while alleviating its main problem by reducing the 'manual' work for data subjects. An example of such a technology is the 'Web 2.0 suicide machine'. This was an application that was mainly active around 2010 and "lets you delete all your energy sucking social-networking profiles, kill your fake virtual friends, and completely do away with your Web2.0 alterego"[3]. At the user's request, this application deleted all the user's social contacts and in the case of for example Facebook and Twitter, also all her posts in one go. As such, the technology ties in to the autonomy of the data subject and empowers her by providing her with a technological power tool that allows her to destroy her presence in a particular platform. While this particular application is not very fine-tuned because it offers the subject a relatively all or nothing choice to deal with the mediating technology's storage and retention capacity, we can also imagine that applications can be developed that will allow users to exercise their autonomy in more nuanced manners. Unfortunately, internet giants like Facebook tend to be not amused with the use of such externally offered tools that provide users with more control over their information and are likely to

[2] In this case, one can think of accelerating the aging process the younger the data subject is. The passing of a year has a bigger impact on the identity of a data subject, the younger the data subject is: there is generally a much more substantial difference between an individual when she was one and when she was two, compared to when she was fifty-two and fifty-three. Accelerating the aging time of files based on the age of the data subject would account for the rate with which identity is likely to change. However, the problem of addressing the issues in this manner is that it does not take into account changes that are related to other things than age (job change, marriage, life choices), and is problematic when two data subjects of a different age are presented in one signifying object (e.g., a picture of a father with child).

[3] `http://suicidemachine.org/`, last accessed 01-08-2019.

prohibit the use of such applications on their platform.[4]

While the above mentioned technical solutions are not perfect (yet) and have their own set of risks, I argue that it is valuable to explore how the right to erasure may take shape as a literal 'counter technology', or more particularly as a 'privacy-enhancing technology' (PET), or as suggested by Novotny & Spiekermann (2014), as a 'oblivion-enhancing technology' (OETs). The GDPR itself requires the implementation of data protection by design and by default (art. 25 GDPR). What such data protection by design and default could look like in the technological praxis, has been researched by Colesky *et al.* (2016). In their article *A critical analysis of privacy design strategies*, they identify several 'privacy design strategies' that can help to shape the access and context of information. Although the majority of these strategies seem to be focused on company-internal systems, I find it worthwhile to briefly mention them and point out which of these many be worthwhile to consider to apply on the Web. One of the suggested tactics is minimisation. Minimising the collection of personal information is also one of the demands listed in the GDPR (art. 5(1)(c) GDPR). It entails a selection of the content, whereby the information that is not needed is excluded, stripped or destroyed. This could be a viable tactic for addressing issues with online content that revolve around abundant information. For content that is necessary to retain as-is, there are several other tactics that can help to nuance the presence of information. For instance, we can consider restricting access. Online, this can take the shape of, for instance, requiring users to have an account or to apply a form of geo-blocking. Another tactic that can be used is the separation of information (Colesky *et al.*, 2016). By isolating information collections, or by distributing their content over different locations, we can reduce the risk that information is combined into a more extensive view on a particular individual. This may be usable in some forms (e.g., the use of robot.txt can make a considerable difference), but is likely only viable for certain applications because the Web's hyperlinking affordances promote the opposite of this strategy. The last tactic that is worth mentioning is abstraction (Colesky *et al.*, 2016). If content is summarised or grouped on a more general information level, the focus of the content shifts from particular individuals to more generic information. As such, the informational visibility of specific individuals is reduced. To what extent these tactics, if any, are truly viable strategies to use on the Web is still a matter for research — one that is worthwhile looking into.

In sum, next to, or in combination with art. 17 GDPR, there is a wide array of possible alternatives to address the issues discussed in this study. While some ways to address the issues at hand seem more promising than others, there is no one-size-fits-all solution, just like there is not one problem. The trick is to find in every case the proper means to address the issue. In general, the technological solutions seem on a practical level the most promising because they can relatively easily bypass the problems of art. 17 GDPR's subject dependency, while they

[4]See e.g., Paul McNamara, "Facebook blocks 'Web 2.0 Suicide Machine'", *Computerworld*, 2010. `https://www.computerworld.com/article/2522527/facebook-blocks--web-2-0-suicide-machine-.html`, last accessed 03-07-2019.

touch the heart of the issues that result from a strong impact of technological intentionality. However, the legal instruments have one significant advantage over the human and technological alternatives: these instruments are equipped with safeguards that prevent a disproportional harm to the interests of others than the data subject. Despite the flaws and drawbacks of art. 17 GDPR, I therefore argue that the right to erasure itself is a reasonable means to address a significant part of the issues, and one that is preferable over some of the alternatives discussed in this section. However, instead of looking for alternatives to replace art. 17 GDPR, it may be worthwhile to look for ways to 'upgrade' the use of art. 17 GDPR by combining it with technological applications in such a manner that its greatest weaknesses can be reduced or overcome.

10.5 Limitations and further research

The present study has limitations. Given the debate surrounding art. 17 GDPR and the reasons provided by Reding for its introduction, I have approached this research from a particular angle: the focus is on the technological mediation by the Web on the front end level. The study revolves around the influence of front end internet technologies on the narrative constructions related to one's identity. With this, the study aims to provide a particular piece of the puzzle. This piece helps to shape and give a foothold to other pieces of the puzzle, but at the same time, it will need these other pieces to reach the full picture. It is therefore important to acknowledge the limitations of this study and use them as directions for future research. In this section, I discuss the main limitations of this study's focus and propose how to move forward with further research.

The first main limitation of this study is that with its focus on the technological mediation itself, I have given less attention to both the poles at either side of the mediation, the user and the subject. On both the person of the user and identity construction of the subject much more could be said. In the case analysis, the focus on the technological mediation required me to clarify the 'who', the 'how', and the 'what' in order to sketch a picture of the presence of online content. The question that I did not touch upon, is the 'why' of users: why do users post content online? Only occasionally I briefly touched upon this 'why', and only to the extent that it allowed me to explain the mechanisms of the issue at hand. This does not mean that the 'why' of users is unimportant. On the contrary: without a reason why, users would not engage in online activities.[5] However, I found that the biggest gap that needed to be addressed is what concretely is happening with the information that, once materialised, has a certain semantic autonomy and is as such detached from the intentions of the author. For this reason, I gave the motivation of encoding users, as well as of viewing users, relatively little attention.

[5]The motivation underlying user behaviour has long been a topic of research in social sciences and psychology, and much progress has already been made on this level, especially with regard to social media use (see e.g., Park *et al.*, 2009; Wang *et al.*, 2015; Lee *et al.*, 2015; Chin *et al.*, 2015).

By doing so, I also sidelined the effect of some specific factors that are worthwhile to tie in to the findings of this research in the future: the manner in which the online mediation of information may differ in its impact based on gender, age, cultural background, as well as how the online informational persona affects group identity. For example, I did not venture into the discriminatory implications of racial profiling (cf. Sweeney, 2013). It is important to note that I left the user, as well as the subject, relatively open — which likely results in a closest resemblance of the situation as it would be for privileged white males. Moreover, with the limitation of the research scope to individuals, the impact of the mediating technology for particular identities and groups, especially those that are not privileged, as well as the functionality of art. 17 GDPR to address group identity issues, remained outside the scope of this study. With this basic undefined individual user, I therefore run the risk of unintentionally reaffirming the position of particular privileged individuals. However, because my goal was to provide some clarity and a foothold with regard to how the Web affects our online portrayal, I hope this study can function as a baseline for future research into specific identity cases. With the help of research in social sciences and identity studies, we can further critically assess the impact of technological intentionality on the construction of individual and group narratives.

Furthermore, by focusing on the impact of Web technology, this study sees to a particular piece of the puzzle. The result is that this study does not provide the full overview that is needed to be able to concretely answer questions with regard to what exactly should be erased, when exactly this should be done, and how exactly this should be done. This requires a full picture of the societal, political, and economic interests and views — related to the specifics of the case — that play a role. The answer as to when personal information can be published and publicly disseminated should therefore come from a wider research in combination with societal debate. The framework developed in this study contributes to this by providing a base and context that helps to answer these questions, and fine-tune the answers as well as the chosen form of erasure if called for. The next step on this front is therefore to further research the balance of various individual, societal, political, cultural, and economic interests in relation to the duration and form of information processing with the framework developed in this study as foundation.

What I take to be the other main limitation of this research, is its focus on the front end of the Web. This restriction leaves highly problematic processing that goes on in the back end out of its scope. As such, this study leaves the potentially hazardous implications of processing by e.g., Facebook, Cambridge Analytica, and others, unexplored, as well as the role of art. 17 GDPR in removing data from the back end level and third party systems that access such data via APIs. During this study I found that adding some remarks and basic analyses of this back end processing of personal information did not seem to do justice to the complexity of this processing and its impact on individuals and society as a whole. Analysing the character of the problems brought about by the back end processing of personal information requires a different kind of research and analysis than I performed in this study. As doing both would double the size of this study, I decided to leave the

focus on the front end given the fact that the debate surrounding art. 17 GDPR so far has been centred on the front end of the Web. However, a research into the problems of back end processing and the right to erasure's ability to address these, is crucial. It is a research that I dearly want take up next.

Last, the focus on the legal side of art. 17 GDPR, and not its concrete application in all its possible technological forms, is a limitation of this study that can benefit from further research. At diverse IT and hacker conferences I have been questioning people with various technological backgrounds on their views, suggestions, and ideas on how to apply erasure in different cases. My initial findings are that these forms can vary per discussed technology, can highly change over time with technological developments, and offer an endless range of options and nuances if one allows one's creativity to run wild. The technical possibilities of concrete applications of art. 17 GDPR is a study on its own and requires a significant amount of technological expertise. This would be an interesting and valuable direction for future research in fields like privacy enhancing technologies and value sensitive design.

10.6 Final thoughts

With the analyses in this study I hope to have filled in some blanks surrounding art. 17 GDPR. Moreover, I hope that with the proposed perspective on art. 17 GDPR, I managed to overcome, or at least soften, some of the oppositions that dominate the debate by shifting the angle towards the manner in which the problems are brought about and the role of technology therein. However, the blanks filled in here are only a few steps forward in the many issues that arise due to the increasing intertwining of the online and the offline that is typical for our information age. What happens online, does not stay online, but spills into the offline. The online viral videos we see become a topic of our lunch conversations, we call a friend when we see on Facebook that she is having a rough time, we look up what we can find on our prospective dates, etc. I only expect this to increase the upcoming years. With merely seconds between the access to online signifying objects, we lose many of the nuances and refinements that we have in our offline information flows. While art. 17 GDPR is not the answer to life, the universe, and everything, it can address particular instances where we can find a too strong expression of technological intentionality in our narrative identities. The use of art. 17 GDPR is not the end of the world or the Web as we know it (online information has been erased far before the introduction of art. 17 GDPR), nor is the right a panacea. There are limits to the viability of art. 17 GDPR to address the issues identified in this study. However, art. 17 GDPR is not a stand alone right: it is part of a bigger instrumentarium that encompasses other rights that may provide a solution. Also, we can look into the direction of users as well as technology to come up with alternative options to address the issues at hand. Several solutions have already been proposed, and I expect in the future we will hear of more and novel ways to address the identified issues.

Whatever route we decide to take, they are all lined with questions and challenges. However, if we want to address the problems identified here, see little salvation in art. 17 GDPR for a particular case, and are unable to think of other options, something needs to give: either we need to address one or more of the causal forces constituted by the technological architecture of the Web, or we need to address how we use the technology. There is always the possibility that what we are experiencing here, are merely growth pains. Maybe we will get used to these problems being an aspect of our lives in the increasingly developing Information Age. In time, the Web as socio-technological construction may grow into a new phase and we will get used to its impact; we may start to use the Web differently, or we may accept and find ways to deal with, or simply stop caring about, the problems that this study identified. Yet, this would ruin the variation in our lives and flatten our identity and many of our interactions and relations into a one-size-fits-all existence. Except for a few brave souls, experimentation would be reduced to what is socially acceptable or considered reasonable in contemporary society. This would upturn what our relation to technology *should* be; instead of adapting the technology to our wishes, we adapt ourselves to the direction in which the technology is moving. And even then, given the implications of the Web for our narrative identity and how it portrays us, the problems of a distorted narrative would remain.

If there is a consistent stream of erasure requests deriving from all sorts of people about various kinds of content, we need to consider that we may have some fundamental problems in our current online praxis. These fundamental problems should not be addressed by art. 17 GDPR. Instead, we should rethink our practices and carefully consider the technology we use. We should be the ones saying what is important about us and our lives, not technology. I therefore argue that we — as subjects, users, information controllers and as society — need to think on how we want to deal with online information flows, and how present we want various personal details of the present and the past to be. We need to think about how we want to shape our onli*f*e. Technology should not be the end of *our* story. Bring a towel.

Acknowledgements

I would like to express my gratitude to prof. Ronald Leenes, prof. Bert-Jaap Koops, the Tilburg Institute for Law, Technology, and Society (TILT) at the Tilburg University, Digital Security at the Radboud University, the Privacy and Identity Lab (PI.lab), and SIDN for making this research possible and supporting me during this project. This research was conducted as part of Privacy and Identity Lab (PI.lab) and funded by SIDN.nl (`https://www.sidn.nl`). Last, I want to thank the committee members, prof. mr. dr. M. Hildebrandt, prof. dr. J.V.J. van Hoboken, prof. mr. E.M.L. Moerel, prof. dr. G. Sartor, dr. M.L. Jones, and dr. A.P. Schouten, for their valuable suggestions for the final version of this dissertation. All remaining errors are my own.

Summary

English summary

The main question to which this study provides an answer is: *To what extent is art. 17 GDPR, the right to erasure ('right to be forgotten'), a viable means to address problems for individuals raised by the presentation of online personal information to Web users?* With this study, I aimed to fill what I perceived to be a gap in the debate surrounding art. 17 GDPR: a clear view on the problems that it can address. The research presented here provides us with a sharpened view on when and how to apply art. 17 GDPR. By providing an in-depth view of the how and what of the problems, as well as the respective roles of technology and human agents herein, the analyses can help to contextualise, and in some cases maybe even overcome, potential conflicts between the right to erasure and the other interests involved, most notably those protected under the right to freedom of expression and information. Moreover, by clarifying how the problems come to be, and what the key elements in the various problems are, the analyses provide handholds for a precise and fine-tuned application of erasure in a manner that respects the various interests at stake.

At the heart of this research lies the relation between individuals, their personal information, and the manner in which a mediating technology can present this in a problematic manner. As human beings, in order to account for the sameness and change of our identity over time and to others, we tell a certain story of who we are: we construct a certain plot and thereby give shape to our narrative identity. However, when we externalise personal information by means of technology, this information we materialise gains a certain autonomous existence, separate from its author. It can tell a particular story to audiences about who we are, thereby constructing an exteriorised narrative identity of us. This narrative is affected by the technology that mediates the information: the mediating material transfers some of its affordances and characteristics to the narrative that it carries. Connecting to the work of most notably Verbeek, I discuss that this autonomously existing material form of the technology has a certain directionality in which it establishes a particular relation between people and the world. This directionality is embodied in the concrete material design of the technology and is a material form of 'intentionality' which is always in

a necessary interplay with human intentionality. By being the materialisation of personal information and presenting it to people, the technology presses its intentionality to the narrative it presents — instilling it with some of its own inclination towards sameness or change. By presenting information in a particular manner, technology 'tells' us more than solely the content of the object. While technology does not construct a narrative in the classical sense, it does affect the story that is told. By doing so, the technology affects the role of the narrator, the audience of the narrative, the manner in which audiences can engage with the narrative, as well as the content. The materialised narrative identity thus entails a complex hybrid intentionality in which the impact of the human and the mediating technology can be intertwined in different manners, with various degrees of human and technological intentionality. The problem is that this externalised narrative identity may not be the same story about sameness and change that we tell about ourselves.

In chapters 4 to chapter 9 I trace how the Web can give rise to problematic portrayals of individuals. I show that on the Web, our social interactions and personal representations are often caught up in a battle for attention between different parties, while the mediating technologies imprint their characteristics on our online narrative identities. In this environment, human beings are produced as a subject to particular plotlines that tell a certain story about their past and present identity. Problems with the narrative identity occur on the level of the emplotment, the context of the narration, the narrator, and the audience selection. When users encode personal information online, this information takes shape in a sphere of hybrid intentionality that is co-shaped by the human agent as well as the mediating technology. For human agents, the Web is a challenging sphere to realise the transmission of personal information in accordance with their wishes: as the Web's technology is the medium and constitutes the environment in which user encodes information, it impresses its own affordances on the processes of encoding, storage, and retrieval. Everyone can publish personal information online, while the information is easily copied, edited, and transmitted. Even more, parts of the online environment are an accelerated and opaque playing field where informational distances are renegotiated in schemes of interest, popularity and profit — which can make it difficult for users to share information only in intended manners. Internet giants like Google and Facebook control important parts of the technology that constitutes the online environment. As these parties have a significant interest in generating user attention for profit, they design the online environment they control in ways that play into their interests. In order to motivate users to spend attention, engage with the technology and provide content, they manage information flows in which meaning is often tied to popularity and advertisement revenue, while they keep the underlying mechanisms of the technology hidden from users.

With the Web's publishing and transmission affordances and the control over the mediating technologies in the hands of various players, the referent's control over the *who* with whom information is shared, the *when*, the *how*, and *the extent* to which this information is shared can all be challenged. The signifying objects

shaped in this realm of various constructions of hybrid intentionality can easily be constructed and reassembled in such a manner that the presented persona becomes an exaggerated, distorted and problematic reflection of the current selves of the referents. Online technologies like search engines can actively affect the emplotment of the materialised narrative. The broad informational scope of the Web combined with centralised information flows can easily lead to the construction of a relatively 'flat character': differentiations between time, space, audiences, context, relations, and public and private are diminished and replaced by differentiations based on association, interest and/or popularity. Meanwhile, the impact of the online persona on our offline lives is intensified as our current Zeitgeist is characterised by being an information society which heavily uses ICT for virtually all aspects of life. The consequence is that our (potentially problematic) online persona plays an important representative role in many parts of our lives: the high frequency with which we engage with the Web leads increasingly to a setting where we *are* our online representations in interactions.

The conclusion of the problem analyses, is that there is not one problem. Instead, there is a group of problems that have a certain family resemblance in the fact that they influence the online materialised narrative identity beyond the wishes and/or expectations of the subject. The 'long-lasting memory' of the Web, the kickstarter of the wish for a 'right to be forgotten', is just one of the issues. A more important generic trait in this family of problems is the Web's connectivity in combination with the affordances of digital objects: content is easily created, edited, and spread beyond expectations. Online personal information can lead to diverse representative problems for the subject, and vary in the factors that play a main role in the manner in which they come to be. The problems revolve to a great degree around the proportionality of a particular reference in relation to the plot of the narrative identity as the subject believes she should be represented. For online content to cause complications, it does not even need to have an extraordinary or negative character: even mundane information can cause a problematic misrepresentation of the identity of an individual.

Overall, the mediating technology plays a prominent role in many of the issues that this study identified. However, it is not just the technological mediation that plays a major role in this: the human agent is often an accomplice to the problems. While humans who process information online are always expressing a hybrid intentionality, their decisions play an important role in this processing. Online, people are encoding new content, as well as copying, editing, and remixing existing content. Such creation of new content is often simplified by online service industries that offer options for the encoding of information with the use of WYSIWYG website generators like Wordpress, and social media like Instagram, Facebook and Twitter. In these cases, the encoding and dissemination of content is industrialised to a greater or lesser degree. Despite the role of the mediating machinery and its often limited transparency and control options, users are the ones that make the final choice in the encoding and dissemination of this content. Human encoded content that itself is experienced as problematic by the referent (these range from singular encodings to a viral outburst), is therefore a problem that for an important

part can be attributed to human intentionality.

The last phase of this study is to assess whether art. 17 GDPR is capable of addressing the identified issues. However, this is not a simple case of applying the article. In chapter 8, I discuss that the right is caught up with some challenges of its own. The main problem of the right is that it seems to suffer from a lack of clarity that could at least partially be attributed to the right's double naming and framing. While 'forgetting' in relation to art. 17 GDPR can serve as a useful concept, the conceptualisation of art. 17 GDPR as a 'right to be forgotten' overall seems to steer our view in suboptimal directions. It clouds the impact of the Web on other factors that affect the presence of personal information, such as space and proportionality, while its conceptualisation is highly metaphorical. I therefore argue that it is better to take the mechanisms of the right as a point of departure for our further investigation. The mechanisms of the article express a particular functionality: they are focused on giving the individual a certain degree of control over her personal information by means of erasure. However, even 'erasure' in the context of art. 17 GDPR seems to have a somewhat metaphorical character: 'erasure' can take on diverse forms, some of which are technically not even forms of erasure, but of blocking. Erasure could entail a full or partial erasure of the object itself, an erasure of an object from view for (particular) users, or by implementing certain manners of processing that reduce the presence of a particular reference. In spite of the possible vagueness that may come with such a metaphorical understanding of erasure, it is exactly by allowing this broad scale of various forms of erasure, that the right can offer a lot of room to manoeuvre in order to resolve the problems while doing justice to the various interests involved.

In chapter 9, I delve further into the fact that art. 17 GDPR is not an absolute right that gives the individual full control over all the processed information relating to her. Instead, the interests of the subject need to be carefully balanced against the interest of the controller, a potential original publisher, and the general public in the ongoing processing of the information. I argue that in order to come to the most balanced solution, we should fine-tune the form of erasure specifically *to the impact of a particular technological mediation* on the produced narrative. Advancing this perspective further, I argue that art. 17 GDPR should help data subjects to reconfigure the technologically mediated narrative so that the subject's freedom to construct her own narrative identity is protected against unreasonable constraints raised by the processing of personal information. Art. 17 GDPR should therefore be aimed at reducing the presence of a reference according to its accuracy and proportionality viewed in relation to the manner in which the narrative is affected on the level of the narrator, the plot and/or the composition of audiences in a manner unwanted by the data subject. As such, it can help individuals to reconfigure their materialised narratives when these are disproportionally shaped by a technological intentionality or at the hands of others who have no preponderant legitimate interest to tell these stories about us. Given the important role of the mediating technology in this, art. 17 GDPR should ideally have a strong focus on the impact of the technological mediation on the narrative identity. Art. 17 GDPR can be highly valuable if we understand it

as being able to function as *counter* technique. With the focus on technological mediation, ideally the right to erasure should have the upper hand in cases where the technological intentionality shapes the narratives beyond human storytelling, intentions and expectations. This would especially be the case when the mediating technology increasingly takes on the role of external narrator, and suggests to offer a 'narrative' of a subject's life. By taking human intentions as an important guiding principle in assessing whether a certain technologically mediated representation of online information should be addressed, this approach places the focus on the value of *human* autonomy. The open interpretation of erasure allows us to look for an application of erasure that is proportional to the interests of others, and that also alleviates the problems for the subject.

Art. 17 GDPR is not able to address all problems in a satisfying manner: the right's dependency on the subject to invoke it may not in all cases be powerful enough to stand up to the technological affordances of the online environment. Additionally, the erasure mechanisms are unable to address in a satisfying manner those issues that revolve around information reaching unintended audiences.

Despite its inability to resolve all the issues, I conclude that overall, art. 17 GDPR is a viable means to address a fair share of the identified problems. It can do well in handling issues that result from a particular persistent presence of a relatively stationary reference with a low quantitative presence in the public sphere — issues that generally arise on basic websites or in search engines. Of these, the ability of art. 17 GDPR to address content in search engines is its biggest asset. Of the websites and online applications that I discuss, search engines have the biggest impact on our narrative identity and generally the strongest expression of technological intentionality therein. Meanwhile, search engines are also the most successful point to apply art. 17 GDPR in order to address problematic narratives spread over the Web. In guiding users towards content, search engines can increase or decrease the chance that a certain user comes in contact with a specific reference. They thus have a powerful influence on the formation of audiences for particular content and can reduce the presence of certain online references. The strong position and gatekeeping role of search engines therefore imbues them with the power to address a fair share of the issues, or at least address the issue in its most intense form.

In the guise of a counter technique, art. 17 GDPR can mitigate the problematic imprints of technological intentionality on our narrative identity, while respecting relevant human intentions with an eye on its value for the controller and the public interest. However, as a legal instrument, art. 17 GDPR is a counter technique, but not a counter *technology* — a practical technological tool. It is thus a different type of instrument than the technological information processing that it aims to address. Although the right to erasure helps to create a better power balance between data subjects vis-à-vis the information controllers on the legal level, it does this in a milieu in which these controllers control and generally better oversee mediating technology. A technical power imbalance between data subject and controllers thus remains. This is tricky because the mediating technology plays a key role in many of the issues that are raised by the presentation of online personal

information to Web users. If we see an ongoing prevalence of erasure requests over time, this may therefore indicate that we need to address this power imbalance at a deeper level than is offered by art. 17 GDPR.

Nederlandse samenvatting

Aan dit onderzoek ligt de volgende hoofdvraag ten grondslag: *In hoeverre is artikel 17 van de Algemene Verordening Gegevensbescherming (AVG), het recht op gegevenswissing ('recht op vergetelheid'), een bruikbaar middel om problemen aan te pakken die voor individuen kunnen ontstaan met de presentatie van online persoonlijke informatie aan webgebruikers?* In het debat rond art. 17 AVG is het hoe en het wat van de problemen die het recht het hoofd zou moeten bieden relatief gezien onderbelicht gebleven. Dit is een gemis, omdat de manier waarop de problemen in elkaar steken van groot belang is voor de vraag of art. 17 AVG überhaupt in staat is om ze te adresseren, alsmede voor de manier waarop art. 17 AVG idealiter toegepast wordt. Met dit onderzoek wil ik de leemte in het debat opvullen door een duidelijk beeld te geven welke problemen art. 17 AVG kan adresseren en hoe het dit op een gebalanceerde manier kan doen. Door inzicht te geven in de totstandkoming van de problemen en de belangrijkste elementen hierin, bieden de analyses in deze studie houvast voor hoe we in de praktijk het recht op gegevenswissing kunnen toepassen op een manier die zo goed mogelijk rekening houdt met de belangen van de verschillende partijen. Dit geldt met name voor de belangen die worden beschermd door het recht op vrijheid van meningsuiting en informatie.

De kern van dit onderzoek gaat in op de relatie tussen mensen, hun persoonlijke informatie en de manier waarop een technologie deze kan presenteren aan gebruikers. Hierin kunnen problemen ontstaan als het door de technologie gepresenteerde beeld afwijkt van hoe individuen hun eigen identiteit ervaren en aan anderen zouden willen presenteren. Van belang hierbij is dat identiteit niet onveranderlijk is, maar zich ontwikkelt, zowel over de tijd heen als in samenspel met anderen. Om rekenschap af te leggen over dergelijke veranderingen in onze identiteit vertellen we als mens een bepaald verhaal over wat ons karakter is: we construeren een bepaalde plot en geven zo vorm aan onze narratieve identiteit. Wanneer we echter informatie over onze identiteit materialiseren in een technologie, krijgt deze informatie een zeker autonoom bestaan, los van de auteur. Deze informatie kan hiermee, los van ons, een bepaald verhaal vertellen aan het publiek over wie we zijn en vormt hiermee een gematerialiseerde versie van onze narratieve identiteit. Aansluitend bij het werk van met name Verbeek, behandel ik hoe deze autonoom bestaande materiële vorm van de technologie een bepaalde gerichtheid heeft in de manier waarop zij een relatie tussen mensen en de wereld tot stand brengt. Deze gerichtheid wordt belichaamd in het concrete materiaalontwerp van de technologie en is een materiële vorm van 'intentionaliteit' die altijd in een noodzakelijk samenspel zit met menselijke intentionaliteit. Door de persoonlijke informatie te materialiseren en deze aan mensen te presenteren, drukt

de technologie haar intentionaliteit op het verhaal dat zij presenteert en geeft het hiermee een zekere neiging tot gelijkheid en/of verandering mee. De technologie beïnvloedt het verhaal dat wordt verteld door informatie op een bepaalde manier te presenteren en een deel van haar eigenschappen over te dragen op de informatie. Hiermee beïnvloedt de technologie de rol van de verteller, de inhoud, de reikwijdte van het publiek, alsmede de manier waarop het publiek kan omgaan met het verhaal. Op deze wijze neemt de technologie deel aan de constructie van de gematerialiseerde narratieve identiteit. De mate waarin de technologie deze beïnvloedt is afhankelijk van zowel de technologie als de mensen die haar gebruiken: het betreft hier een complexe hybride intentionaliteit, waarbij de impact van mens en technologie op verschillende manieren met elkaar verweven kunnen worden in verschillende gradaties van menselijke en technologische intentionaliteit. De resulterende gematerialiseerde narratieve identiteit kan hierdoor een ander verhaal vertellen over het karakter van het subject dan het over zichzelf zou vertelt.

In de hoofdstukken 4 tot en met 9 onderzoek ik hoe persoonlijke informatie op het Web een individu op een problematische wijze kan representeren. In de analyses laat ik zien dat op het Web onze sociale interacties en persoonlijke representaties gemakkelijk verstrikt raken in een strijd om aandacht tussen verschillende partijen, terwijl de mediërende technologieën hun karakteristieken inprenten op de gepresenteerde narratieve identiteiten. In deze omgeving worden mensen geproduceerd als het subject van verhaallijnen die een beeld geven van hun identiteit. Problemen met de wijze waarop mensen gerepresenteerd worden doen zich voor op het niveau van de verhaallijn, de context van het verhaal, de verteller en de selectie van het publiek. Vanwege de vaak meervoudige hybride intentionaliteit die wordt gevormd door het samenspel van menselijke actoren en technologie, is het voor gebruikers lastig om via het Web informatie over te dragen op een manier die volledig in overeenstemming is met hun wensen: het Web is een complex medium waarin iedereen persoonlijke informatie kan publiceren, kopieëren, bewerken, aan andere inhoud linken, en verder kan doorsturen. Dit wordt extra bemoeilijkt wanneer mensen hun informatie delen in een technologische omgeving die in handen is van derde partijen, en waar zij zelf maar een gebrekkige controle en beperkt overzicht hebben over de gebruikte technologie. Internetgiganten zoals Google en Facebook beheersen belangrijke delen van de technologie waardoor de online omgeving wordt gevormd. Deze partijen hebben een aanzienlijk belang bij het genereren van aandacht van gebruikers voor winst en hebben de online omgeving die zij beheren, vormgegeven in overeenstemming met hun belangen. Terwijl ze proberen veelal de aandacht van gebruikers vast te houden en ze te verleiden om meer informatie te delen, beheren dergelijke spelers informatiestromen waarin betekenis vaak gekoppeld is aan populariteit en advertentie-inkomsten, terwijl ze de onderliggende mechanismen van de technologie voor gebruikers verborgen houden. Webgebruikers bevinden zich zodoende vaak in een ondoorzichtig speelveld waarin de toegankelijkheid en context van informatie wordt bepaald door interesses, populariteit en winst. Dit bemoeilijkt voor hen het gecontroleerd delen van informatie.

Door de publicatie- en transmissiemogelijkheden van het Web en met de

technologieën in handen van verschillende actoren heeft het individu hooguit een beperkte controle over met *wie*, *wanneer*, *hoe* en *de mate* waarin haar informatie wordt gedeeld. In deze omgeving van hybride intentionaliteit kan het gepresenteerde door verschillende actoren en langs verschillende wegen narratief gevormd worden. Online kan gemakkelijk een overdreven, vervormd en/of onjuist beeld van een individu ontstaan. Technologieën zoals zoekmachines kunnen hierin een vergaande rol spelen. Offline differentiaties tussen tijd, ruimte, doelgroepen, context, en relaties worden vervangen door differentiaties op basis van associatie, interesse en/of populariteit. Hierin wordt het subject gemakkelijk geproduceerd als een 'vlak' karakter. Ondertussen kan het beeld dat online van ons gepresenteerd wordt, een grote impact op ons hebben omdat onze huidige maatschappij gekenmerkt wordt door een intensief gebruik van ICT voor vrijwel alle aspecten van het leven. Het gevolg is dat de online presentatie van onze identiteit in veel aspecten van het leven een belangrijke representatieve rol speelt: in veel interacties *zijn* wij in feite onze online representatie.

De conclusie van de probleemanalyses is dat er niet één probleem is, maar meerdere problemen die een zekere familiegelijkenis vertonen in het feit dat ze de online gematerialiseerde narratieve identiteit beïnvloeden buiten menselijke wensen en/of verwachtingen om. Het zogenaamde 'ijzeren geheugen' van het Web, de aanleiding van de wens om een 'recht op vergetelheid' in het leven te roepen, is slechts een van die problemen. Een belangrijkere onderliggende eigenschap in deze familie van problemen is de connectiviteit van het Web in combinatie met de mogelijkheden die digitale objecten bieden: informatie wordt eenvoudig gecreëerd, op het Web geplaatst, bewerkt en verspreid. De problemen draaien in grote mate om de evenredigheid van een bepaalde referentie in relatie tot de verhaallijn van de narratieve identiteit van het individu. Online informatie hoeft niet bijzonder of negatief te zijn om complicaties te veroorzaken: ook een alledaags element dat als representatief voor een individu gepresenteerd wordt, kan al aanleiding zijn voor misvattingen.

Over het algemeen speelt de bemiddelende technologie een prominente rol in veel van de onderzochte problemen. Het is echter niet alleen de technologie die een belangrijke rol speelt: de mens is vaak medeplichtig aan de problemen. Ofschoon mensen die informatie online verwerken, altijd uiting geven aan een hybride intentionaliteit, spelen hun beslissingen wel degelijk een belangrijke rol. Mensen kunnen nieuwe informatie op het Web plaatsen en bestaande informatie kopiëren, bewerken en verder verspreiden. Deze vormen van informatieverwerking worden vaak vereenvoudigd aangeboden in online applicaties zoals Wordpress en in sociale media zoals Instagram, Facebook en Twitter. In deze gevallen wordt het uploaden, bewerken en verspreiden van informatie in meer of mindere mate geïndustrialiseerd. Ondanks de rol van de technologie hierin, zijn het de gebruikers die er uiteindelijk voor kiezen om bepaalde informatie te verwerken. Informatie die op zichzelf als problematisch wordt ervaren, is daarom voor een groot deel toe te wijten aan menselijke intentionaliteit.

De laatste fase van dit onderzoek bestaat uit het toetsen in hoeverre art. 17 AVG de geïdentificeerde problemen kan oplossen. Hiervoor kan het artikel

echter niet eenvoudigweg worden toegepast op de casussen: het artikel zelf kampt met de nodige problemen en onduidelijkheden. Dit leg ik uit in hoofdstuk 8. Een deel van deze problemen en onduidelijkheden hangt samen met de dubbele naamgeving van het recht. Hoewel 'vergeten' in relatie tot art. 17 AVG een nuttige metafoor kan zijn, lijkt de conceptualisering van art. 17 AVG als een 'recht op vergetelheid' ons over het algemeen in een suboptimale richting te sturen. Het vertroebelt het zicht op de impact van het Web op andere factoren die de aanwezigheid van persoonlijke informatie beïnvloeden, zoals ruimte, context en proportionaliteit. De mechanismen van het recht vormen daarom een beter uitgangspunt voor het verdere onderzoek. Deze mechanismen drukken een bepaalde functionaliteit uit: ze zijn gericht op het geven van een zekere mate van controle aan individuen over hun persoonlijke informatie door mogelijkheid tot het wissen van gegevens. Zelfs 'wissen' in de context van art. 17 AVG lijkt echter een enigszins metaforisch karakter te hebben: 'wissen' kan verschillende vormen aannemen, waarvan sommige technisch gezien niet eens vormen van wissen zijn, maar van blokkeren. Ondanks de mogelijke onduidelijkheid die gepaard kan gaan met zo'n metaforisch begrip van wissen, is het juist door verschillende vormen van wissen mogelijk te maken dat het recht veel ruimte biedt om de problemen op te lossen op een manier die recht kan doen aan de belangen van de verschillende betrokkenen.

In hoofdstuk 9 ga ik dieper in op het feit dat art. 17 AVG geen absoluut recht is dat individuen de volledige controle geeft over alle informatie die naar hen verwijst. In plaats daarvan moeten de belangen van het individu zorgvuldig worden afgewogen tegen de belangen van de verwerkingsverantwoordelijke, een eventuele originele publiceerder, en het algemene publiek. Om tot een evenwichtige oplossing te komen, is het vab belang dat de vorm van wissen afgestemd wordt op de impact die de bemiddelende technologie heeft op het geproduceerde narratief. Art. 17 AVG moet de aanwezigheid van een referentie proportioneel verminderen op basis van de manier waarop deze het narratief ongewenst beïnvloedt op het niveau van de verteller, de plot en/of de samenstelling van het publiek. Hiermee kan art. 17 AVG individuen helpen om hun online gematerialiseerde narratieve identiteit te herconfigureren wanneer deze onevenredig wordt gevormd onder de invloed van technologische intentionaliteit of door toedoen van anderen die geen zwaarwegend legitiem belang hebben om deze verhalen over ons te vertellen. Gezien de belangrijke rol van technologie hierin zou art. 17 AVG idealiter een sterke focus moeten hebben op de impact van de technologie op de gematerialiseerde narratieve identiteit. In deze context kan het daarom waardevol zijn om art. 17 AVG op te vatten als een recht dat kan functioneren als *contra*-techniek. Door de toepassing van art. 17 AVG te verfijnen in de context van de impact van de technologie, kan dit artikel recht doen aan de belangen van betrokkenen door de mogelijkheid tot een verfijnde balans te creëen om specifieke problemen te adresseren. De open interpretatie van wissen stelt ons in staat om te zoeken naar een toepassing van wissen die evenredig is aan de belangen van anderen, maar die tevens de problemen voor het individu aanpakt of in ieder geval vermindert. Met de nadruk op de impact van technologie zou idealiter het recht op gegevenswissing de overhand

moeten hebben in gevallen waarin de technologische intentionaliteit de verhalen vormt buiten menselijke intenties en verwachtingen om. Dit geldt zeker in het geval van een technologie die in toenemende mate de rol van verteller op zich neemt en 'eigenhandig' een narratief over het individu construeert. Door menselijke intenties als een belangrijk leidend principe te nemen bij de beoordeling of een bepaalde technologisch gemedieerde weergave van online informatie moet worden aangepakt, legt deze benadering de nadruk op de waarde van *menselijke* autonomie. Door menselijke intenties centraal te stellen en vooral een te sterke impact van technologische intentionaliteit aan te pakken, sluit deze conceptualisering van art. 17 AVG aan bij de grondgedachte van de AVG: "de verwerking van persoonsgegevens moet ten dienste van de mens staan" (overweging 4 AVG). Het nemen van menselijke belangen als leidraad sluit tevens aan bij de algemene focus van de AVG op de doelen van de verwerkingsverantwoordelijke: door de legitimiteit van de verwerking te koppelen aan de doeleinden van diezelfde verwerkingsverantwoordelijke spelen haar intenties met betrekking tot de verwerking van de informatie een belangrijke rol in de belangenafweging.

Art. 17 AVG kan niet alle problemen oplossen: het recht is sterk afhankelijk van de menselijke vaardigheden van het individu dat het recht inroept, en hierdoor is het niet altijd krachtig genoeg om de impact van technologisch gedreven informatieverwerking het hoofd te bieden. Bovendien is het wissen van informatie een onbevredigende oplossing in gevallen waarin men de informatie alleen met beperkte publieken wil delen.

Ondanks het feit dat art. 17 AVG niet geschikt is om alle problemen op te lossen, concludeer ik dat het in ieder geval wel in staat lijkt om een redelijk deel van de problemen te adresseren. Art. 17 AVG is het best toegerust om problemen aan te pakken die het gevolg zijn van een bepaalde, relatief stationaire referentie met een langdurige, wellicht kwalitatief prominente, maar kwantitatief lage aanwezigheid in de publieke sfeer. Het problematische plot gecreëerd door de aanwezigheid van deze referentie kan opnieuw worden geconfigureerd door (delen van) de inhoud te wissen, de identificeerbare aspecten ervan te verbergen, of door de toegankelijkheid en/of zichtbaarheid van het object te verminderen. Met deze hoofdfunctionaliteit is art. 17 AVG het meest geschikt om problemen aan te pakken die zich voordoen op reguliere webpagina's of in zoekmachines. Vooral de mogelijkheid tot het verwijderen van persoonlijke informatie uit de informatieverstrekking door zoekmachines is hierbij van belang. Zoekmachines hebben een grote impact op de presentatie van onze online narratieve identiteit én drukken hierin een sterke technologische intentionaliteit uit. Daarnaast kan door het verwijderen van zoekresultaten ook de impact verminderd worden problematische narratieven die verspreid zijn over het Web, omdat de kans dat gebruikers met die informatie in contact komen, hiermee verkleind wordt. Het verwijderen van zoekresultaten kan zelfs een deel van de problematische impact van een virale informatiestroom verminderen.

Hoewel art. 17 AVG in zijn vorm als contra-techniek in staat is om een significant deel van de problemen op een gebalanceerde wijze te adresseren, is het belangrijk te benadrukken dat art. 17 AVG geen echte contra-technologie

is. Het is een juridisch instrument en heeft hierdoor een ander karakter dan de technologische informatieverwerking die het beoogt aan te pakken. Hoewel het recht op gegevenswissing helpt bij het creëren van een betere balans tussen betrokkenen tegenover de verwerkersverantwoordelijken op juridisch niveau, doet het dit in een omgeving waarin deze verwerkingsverantwoordelijken meester zijn over de technologie. Er blijft dus een zekere technische machtsongelijkheid bestaan. Dit is een heikel punt, omdat de technologie een sleutelrol speelt in veel van de problemen die worden opgeworpen door de presentatie van online beschikbare persoonlijke informatie aan webgebruikers. Als we in de loop der tijd een steeds groter wordende stroom aan verwijderingsverzoeken zien, kan dit erop wijzen dat we deze machtsongelijkheid op een dieper niveau moeten aanpakken dan mogelijk is met art. 17 AVG.

Bibliography

Aamodt, Agnar, & Nygård, Mads. 1995. Different roles and mutual dependencies of data, information, and knowledge—an AI perspective on their integration. *Data & Knowledge Engineering*, **16**(3), 191–222.

Accapadi, Jos M, & Weber, Lynne M. 2011 (Nov. 17). *Aging file system.* US Patent App. 12/779,982.

Ackoff, Russell L. 1989. From data to wisdom. *Journal of applied systems analysis*, **16**(1), 3–9.

Acquisti, Alessandro, Brandimarte, Laura, & Loewenstein, George. 2015. Privacy and human behavior in the age of information. *Science*, **347**(6221), 509–514.

Adams, Douglas. 1996. *The Ultimate Hitchhiker's Guide.* Wing Books.

Agamben, Giorgio. 2009. *"What is an apparatus?" and other essays.* Stanford University Press.

Agre, Philip E. 1997. Introduction. *Chap. 1, pages 29–62 of:* Agre, Philip E., & Rotenberg, Marc (eds), *Technology and privacy: The new landscape.* Mit Press.

Ambrose, Meg Leta. 2012. It's about time: Privacy, information lifecycles, and the right to be forgotten. *Stanford Technology Law Review*, **16**.

Ambrose, Meg Leta. 2013. *Digital Oblivion: The Right to be Forgotten in the Internet Age.* ATLAS Institute Graduate Theses & Dissertations 2.

Ambrose, Meg Leta, & Ausloos, Jef. 2013. The right to be forgotten across the pond. *Journal of Information Policy*, **3**, 1–23.

Ausloos, J. 2018. *The Right to Erasure: Safeguard for Informational Self-Determination in a Digital Society?* KU Leuven.

Ausloos, Jef. 2012. The 'right to be forgotten'–worth remembering? *Computer law & security review*, **28**(2), 143–152.

Baddeley, A, Eysenck, Michael W, & Anderson, MC. 2009. *Memory. London.* Psychology Press.

Baker, Paul, & Potts, Amanda. 2013. 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies*, **10**(2), 187–204.

Bar-Tura, Asaf. 2010. Wall-to-Wall or Face-to-Face. *Chap. 20, pages 231–239 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind.* Chicago and La Salle: Open Court.

Barad, Karen. 2007. *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning.* Duke University Press.

Barasch, Alixandra, & Berger, Jonah. 2014. Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research*, **51**(3), 286–299.

Barnes, Susan B. 2006. A privacy paradox: Social networking in the United States. *First Monday*, **11**(9).

Barnet, Belinda. 2013. *Memory machines: The evolution of hypertext.* Anthem Press.

Baroncelli, Lauane, & Freitas, Andre. 2011. *The visibility of the self on the Web: A struggle for recognition.* `http://www.websci11.org/fileadmin/websci/Posters/191_paper.pdf`. Last accessed: 09-10-2018.

Bartolini, Cesare, & Siry, Lawrence. 2016. The right to be forgotten in the light of the consent of the data subject. *Computer Law & Security Review*, **32**(2), 218–237.

Baudrillard, Jean. 1994. *Simulacra and simulation.* University of Michigan press.

Baym, Nancy K, & boyd, danah. 2012. Socially mediated publicness: an introduction. *Journal of Broadcasting & Electronic Media*, **56**(3), 320–329.

Beer, David. 2009. Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, **11**(6), 985–1002.

Benjamin, Walter. 2008. *The work of art in the age of mechanical reproduction.* Vol. 10. Penguin London.

Bennett, Steven C. 2012. The right to be forgotten: Reconciling EU and US perspectives. *Berkeley J. Int'l L.*, **30**, 161.

Berardi, Franco. 2009a. *Precarious rhapsody: Semiocapitalism and the pathologies of the post-alpha generation.* Minor Compositions.

Berardi, Franco. 2009b. *The soul at work: From alienation to autonomy.* Semiotext(e).

Berardi, Franco. 2011. *After the future.* AK Press.

Berger, Jonah. 2013. Beyond viral: Interpersonal communication in the Internet age. *Psychological Inquiry*, **24**(4), 293–296.

Berger, Jonah, & Milkman, Katherine L. 2012. What makes online content viral? *Journal of marketing research*, **49**(2), 192–205.

Berger, Jonah, & Milkman, Katherine L. 2013. Emotion and virality: what makes online content go viral? *GfK Marketing Intelligence Review*, **5**(1), 18–23.

Bernal, Paul. 2011. A Right to Delete? *European Journal of Law and Technology*, **2**(2).

Berners-Lee, T., & Connolly, D. 1995 (Nov.). *Hypertext Markup Language - 2.0*. RFC 1866 (Historic). Obsoleted by RFC 2854.

Berners-Lee, T., Fielding, R., & Masinter, L. 2005 (jan). *Uniform Resource Identifier (URI): Generic Syntax*. RFC 3986 (INTERNET STANDARD). Updated by RFCs 6874, 7320.

Berners-Lee, Tim. 2011. Linked data-design issues (2006). *W3C*. Last accessed: 14-07-2017.

Bjork, Robert A. 1970. Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, **9**(3), 255–268.

Blanchette, Jean-François, & Johnson, Deborah G. 2002. Data retention and the panoptic society: The social benefits of forgetfulness. *The Information Society*, **18**(1), 33–45.

Blood, Rebecca. 2004. How blogging software reshapes the online community. *Communications of the ACM*, **47**(12), 53–55.

Bogost, Ian. 2010. Ian became a fan of Marshall McLuhan on Facebook and suggested you become a fan too. *Chap. 3, pages 21–32 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Boisot, Max, & Canals, Agustí. 2004. Data, information and knowledge: have we got it right? *Journal of Evolutionary Economics*, **14**(1), 43–67.

Bolton III, Robert Lee. 2014. The right to be forgotten: Forced amnesia in a technological age. *J. Marshall J. Info. Tech. & Privacy L.*, **31**, 132.

Bougiakiotis, Emmanouil. 2016. The enforcement of the Google Spain ruling. *International Journal of Law and Information Technology*, **24**(4), 311–342.

boyd, danah. 2010. Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. *Chap. 2, pages 39–58 of:* Papacharissi, Zizi (ed), *A networked self: Identity, community, and culture on social network sites*. Routledge.

boyd, danah. 2014. *The Networked Nature of Algorithmic Discrimination*. Ph.D. thesis, Fordham University.

Brandtzæg, Petter Bae, & Heim, Jan. 2009. Why people use social networking sites. *Pages 143–152 of: International Conference on Online Communities and Social Computing*. Springer.

Brin, Sergey, & Page, Lawrence. 2012. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*, **56**(18), 3825–3833.

Brockmeier, Jens. 2002. Remembering and forgetting: Narrative as cultural memory. *Culture & Psychology*, **8**(1), 15–43.

Brouwer, E.R. 2011. *Legality and Data Protection Law: The Forgotten Purpose of Purpose Limitation*. Kluwer Law International. Pages 273–294.

Brown Jr, James J. 2008. Evil Bert Laden: ViRaL Texts, Community, and Collision. *Fast Capitalism*, **4**(1).

Broxton, Tom, Interian, Yannet, Vaver, Jon, & Wattenhofer, Mirjam. 2013. Catching a viral video. *Journal of Intelligent Information Systems*, **40**(2), 241–259.

Bruner, Jerome. 1994. The "remembered" self. *The remembering self: Construction and accuracy in the self-narrative*, 41–54.

Bucher, Taina. 2012. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, **14**(7), 1164–1180.

Buitelaar, JC. 2014. Privacy and narrativity in the internet era. *The Information Society*, **30**(4), 266–281.

Bunn, Anna. 2015. The curious case of the right to be forgotten. *Computer Law & Security Review*, **31**(3), 336–350.

Bunyard, Tom. 2017. *Debord, Time and Spectacle: Hegelian Marxism and Situationist Theory*. Brill.

Burch, Robert. 2018. Charles Sanders Peirce. *In:* Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, winter 2018 edn. Metaphysics Research Lab, Stanford University.

Burgess, Jean. 2008. All your chocolate rain are belong to us? *Viral videos, YouTube and dynamics of participatory culture,(ed) Video Vortex Reader: Responses to YouTube. Institue of Network Cultures, Amsterdam*, 101–109.

Burkell, Jacquelyn Ann. 2016. Remembering me: big data, individual identity, and the psychological necessity of forgetting. *Ethics and Information Technology*, **18**(1), 17–23.

Butcher, Samuel Henry. 1951. *Aristotle's theory of poetry and fine art: with a critical text and translation of the Poetics. With a prefatory essay, Aristotelian literary criticism.* Vol. 42. Courier Corporation.

Calo, Ryan. 2016. Can Americans resist surveillance. *U. Chi. L. Rev.*, **83**, 23.

Campanelli, Vito. 2014. Frictionless Sharing: The Rise of Automatic Criticism. *Society of the Query Reader. Reflections on Web Search, Amsterdam, Institute of Network Cultures*, 41–49.

Campbell, Marilyn A. 2005. Cyber Bullying: An Old Problem in a New Guise? *Australian journal of Guidance and Counselling*, **15**(01), 68–76.

Carr, David. 1991. *Time, narrative, and history*. Indiana University Press.

Carr, David. 2012. Experience and History. *In:* Zahavi, Dan (ed), *The Oxford handbook of contemporary phenomenology*. Oxford University Press Oxford, UK.

Carr, Nicholas. 2010. *The shallows: How the internet is changing the way we think, read and remember*. Atlantic Books Ltd.

Castells, Manuel. 2002. *The Internet galaxy: Reflections on the Internet, business, and society*. Oxford University Press on Demand.

Cha, Meeyoung, Benevenuto, Fabrício, Haddadi, Hamed, & Gummadi, Krishna. 2012. The world of connections and information flow in twitter. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, **42**(4), 991–998.

Chabot, Pascal. 2013. *The philosophy of Simondon: Between technology and individuation*. A&C Black.

Chander, Anupam. 2016. The racist algorithm. *Mich. L. Rev.*, **115**, 1023–1045.

Chapman, Michael, Ostwald, Michael J, Tucker, Chris, & Bromberek, Zbigniew. 2004. Semiotics, Interpretation and Political Resistance in Architecture. *Pages 384–390 of: The 38th International Conference of Architectural Science Association ANZAScA, "Contexts of architecture", Launceston, Tasmania*.

Chin, Chih-Yu, Lu, Hsi-Peng, & Wu, Chao-Ming. 2015. Facebook users' motivation for clicking the "Like" button. *Social Behavior and Personality: an international journal*, **43**(4), 579–592.

Choi, Yoon Hyung, & Bazarova, Natalya N. 2015. Self-disclosure characteristics and motivations in social media: Extending the functional model to multiple social network sites. *Human Communication Research*, **41**(4), 480–500.

Clark, Andy. 2003. *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford University Press.

Clarke, Roger. 1994. The digital persona and its application to data surveillance. *The information society*, **10**(2), 77–92.

Coeckelbergh, Mark, & Reijers, Wessel. 2016. Narrative technologies: A philosophical investigation of the narrative capacities of technologies by using Ricoeur's narrative theory. *Human Studies*, **39**(3), 325–346.

Cofone, Ignacio. 2015. Google v. Spain: A Right to Be Forgotten. *Chi.-Kent J. Int'l & Comp. L.*, **15**, 1.

Cohen, Nicole S. 2013. Labor Online: Social Media, Audiences, and Advertising. *The Routledge companion to advertising and promotional culture*, 177.

Colesky, Michael, Hoepman, Jaap-Henk, & Hillen, Christiaan. 2016. A critical analysis of privacy design strategies. *Pages 33–40 of: 2016 IEEE Security and Privacy Workshops (SPW)*. IEEE.

Commodore Business Machines, Inc. 1982. *Commodore 64 user's guide*. Commodore 64 Business Machines, Inc. and Howard W. Sams & Co., Inc.

Condella, Craig. 2010. Why Can't We Be Virtual Friends? *Chap. 10, pages 111–121 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Cormode, Graham, & Krishnamurthy, Balachander. 2008. Key differences between Web 1.0 and Web 2.0. *First Monday*, **13**(6).

Cunningham, McKay. 2017. Privacy law that does not protect privacy, forgetting the right to be forgotten. *Buffalo Law Review*, **65**, 495.

Cuonzo, M.A. 2010. Gossip and the evolution of facebook. *Chap. 15, pages 173–179 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind.* Chicago and La Salle: Open Court.

Dafonte-Gomez, Alberto. 2015. The key elements of viral advertising. From motivation to emotion in the most shared videos. *arXiv preprint arXiv:1505.02002.*

Daigle, L. 2004 (sep). *WHOIS Protocol Specification.* RFC 3912 (Draft Standard).

de Andrade, Norberto Nuno Gomes. 2014. Oblivion: the right to be different... from oneself: re-proposing the right to be forgotten. *Pages 65–81 of: The Ethics of Memory in a Digital Age.* Springer.

De Baets, Antoon. 2016. A historian's view on the right to be forgotten. *International Review of Law, Computers & Technology*, **30**(1-2), 57–66.

de Fina, Anna. 2016. Storytelling and audience reactions in social media. *Language in Society*, **45**(4), 473–498.

De Hert, Paul, & Czerniawski, Michal. 2016. Expanding the European data protection scope beyond territory: Article 3 of the General Data Protection Regulation in its wider context. *International Data Privacy Law*, **6**(3), 230–243.

De Hert, Paul, & Papakonstantinou, Vagelis. 2015. Google Spain: Addressing critiques and misunderstandings one year later. *Maastricht Journal of European and Comparative Law*, **22**(4), 624–638.

de Mars, Sylvia, & O'Callaghan, Patrick. 2016. Privacy and Search Engines: Forgetting or Contextualizing? *Journal of Law and Society*, **43**(2), 257–284.

de Meo, Pasquale, Ferrara, Emilio, Fiumara, Giacomo, & Provetti, Alessandro. 2012. On Facebook, most ties are weak. *arXiv preprint arXiv:1203.0535.*

De Montjoye, Yves-Alexandre, Hidalgo, César A, Verleysen, Michel, & Blondel, Vincent D. 2013. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, **3**, 1376.

de Mul, Jos. 2002. *Cyberspace odyssee.* Klement.

de Ridder, Isabelle. 2002. Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process. *Language Learning & Technology*, **6**(1), 123–146.

de Terwangne, Cécile. 2014. The Right to be Forgotten and Informational Autonomy in the Digital Environment. *Pages 82–101 of: The Ethics of Memory in a Digital Age.* Springer.

de Vries, Katja. 2010. Identity, profiling algorithms and a world of ambient intelligence. *Ethics and information technology*, **12**(1), 71–85.

Debord, Guy. 1977. *Society of the Spectacle. rev. ed.* Detroit: Black & Red.

Deleuze, Gilles. 1992. Postscript on the Societies of Control. *October*, **59**, 3–7.

Dennis, Kingsley. 2008. Keeping a close watch–the rise of self-surveillance and the threat of digital exposure. *The Sociological Review*, **56**(3), 347–357.

Derks, Daantje, & Bakker, Arnold B. 2014. Smartphone use, work–home interference, and burnout: A diary study on the role of recovery. *Applied Psychology*, **63**(3), 411–440.

Derrida, Jacques. 1981. Plato's Pharmacy. *Pages 61–171 of:* Johnson, Barbara (ed), *Dissemination.* Chicago, IL: University of Chicago Press.

Deterding, Sebastian. 2012. Gamification: designing for motivation. *ACM Interactions*, **19**(4), 14–17.

DeVito, Michael A, Birnholtz, Jeremy, & Hancock, Jeffery T. 2017. Platforms, people, and perception: Using affordances to understand self-presentation on social media. *Pages 740–754 of: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM.

Dhenakaran, SS, & Sambanthan, K Thirugnana. 2011. Web crawler-an overview. *International Journal of Computer Science and Communication*, **2**(1), 265–267.

Dholakia, Nikhilesh, & Zwick, Detlev. 2001. Privacy and consumer agency in the information age: between prying profilers and preening webcams. *Journal of Research for Consumers*.

Diaz, Alejandro. 2008. Through the Google goggles: Sociopolitical bias in search engine design. *Pages 11–34 of: Web search*. Springer.

DiMaggio, Paul, Hargittai, Eszter, Neuman, W Russell, & Robinson, John P. 2001. Social implications of the Internet. *Annual review of sociology*, **27**(1), 307–336.

Dobele, Angela, Lindgreen, Adam, Beverland, Michael, Vanhamme, Joëlle, & van Wijk, Robert. 2007. Why pass on viral messages? Because they connect emotionally. *Business Horizons*, **50**(4), 291–304.

Dodge, Martin, & Kitchin, Rob. 2007. "Outlines of a world coming into existence": Pervasive computing and the ethics of forgetting. *Environment and Planning B*, **34**(3), 431–445.

Dommering, E.J. 2005. Annotatie bij EHRM 24 juni 2004 (Caroline von Hannover / Duitsland). *NJ*, **22**.

Downes, Stephen. 1999. Hacking memes. *First Monday*, **4**(10).

Downey, Greg. 2014. Making media work: Time, space, identity, and labor in the analysis of information and communication infrastructures. *Media technologies: Essays on communication, materiality, and society*, 141–166.

Duntemann, Jeff. 1992. *Assembly language step-by-step*. John Wiley & Sons.

Dyer-Witheford, Nick. 2015. *Cyber-proletariat: Global labour in the digital vortex*. Between the Lines.

Eastman, Jacqueline K, & Iyer, Rajesh. 2005. The impact of cognitive age on Internet use of the elderly: an introduction to the public policy implications. *International Journal of Consumer Studies*, **29**(2), 125–136.

Eckler, Petya, & Bolls, Paul. 2011. Spreading the virus: Emotional tone of viral advertising and its effect on forwarding intentions and attitudes. *Journal of Interactive Advertising*, **11**(2), 1–11.

Edunov, Sergey, Diuk, Carlos, Filiz, Ismail Onur, Bhagat, Smriti, & Burke, Moira. 2016. Three and a half degrees of separation. *Research at Facebook*.

Eggers, Dave. 2013. *The Circle*. McSweeney's.

Elers, Steve. 2014. Maori are scum, stupid, lazy: maori according to Google. *Te Kaharoa*, **7**(1).

Ellis, Katie, & Kent, Mike. 2010. Tweeters take responsibility for an accessible web 2.0. *Fast Capitalism*, **7**(1).

Ellison, Nicole B, & boyd, danah m. 2007. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, **13**(1), 210–230.

Eslami, Motahhare, Rickman, Aimee, Vaccaro, Kristen, Aleyasen, Amirhossein, Vuong, Andy, Karahalios, Karrie, Hamilton, Kevin, & Sandvig, Christian. 2015. I always assumed that I wasn't really that close to [her]: Reasoning about Invisible Algorithms in News Feeds. *Pages 153–162 of: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM.

Fairweather, A, & Halpern, J. 2010. Do status updates have any value. *Chap. 17, pages 191–199 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Falk, R Frank, & Miller, Nancy B. 1998. The reflexive self: A sociological perspective. *Roeper Review*, **20**(3), 150–153.

Falkinger, Josef. 2007. Attention economies. *Journal of Economic Theory*, **133**(1), 266–294.

Fazlioglu, Muge. 2013. Forget me not: the clash of the right to be forgotten and freedom of expression on the Internet. *International Data Privacy Law*, **3**(3), 149–157.

Feenberg, Andrew. 2002. *Transforming technology: A critical theory revisited*. Oxford University Press.

Feenberg, Andrew. 2010. *Between reason and experience: Essays in technology and modernity*. MIT Press.

Ferguson, Douglas A, & Perse, Elizabeth M. 2000. The World Wide Web as a functional alternative to television. *Journal of broadcasting & electronic media*, **44**(2), 155–174.

Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., & Berners-Lee, T. 1999 (jun). *Hypertext Transfer Protocol – HTTP/1.1*. RFC 2616 (Draft Standard). Obsoleted by RFCs 7230, 7231, 7232, 7233, 7234, 7235, updated by RFCs 2817, 5785, 6266, 6585.

Fiesler, Casey, Dye, Michaelanne, Feuston, Jessica L, Hiruncharoenvate, Chaya, Hutto, Clayton J, Morrison, Shannon, Khanipour Roshan, Parisa, Pavalanathan, Umashanthi, Bruckman, Amy S, De Choudhury, Munmun, *et al.* 2017. What (or who) is public?: Privacy settings and social media content sharing. *Pages 567–580 of: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM.

Fiske, John. 2010. *Introduction to communication studies*. Routledge.

Floridi, Luciano. 2005. The ontological interpretation of informational privacy. *Ethics and Information Technology*, **7**(4), 185–200.

Floridi, Luciano. 2009. Against digital ontology. *Synthese*, **168**(1), 151–178.

Floridi, Luciano. 2011. *The Philosophy of Information*. Oxford University Press.

Floridi, Luciano. 2014. *The fourth revolution: How the infosphere is reshaping human reality*. OUP Oxford.

Floridi, Luciano. 2015. *The onlife manifesto*. Springer.

Floridi, Luciano. 2016. Semantic Conceptions of Information. *In:* Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, spring 2016 edn.

Flusser, Vilém. 1990. On memory (electronic or otherwise). *Leonardo*, 397–399.

Flusser, Vilém. 2011. *Into the universe of technical images*. Vol. 32. University of Minnesota Press.

Fogg, Brian J. 1999. Persuasive technologies. *Communications of the ACM*, **42**(5), 27–29.

Frantziou, Eleni. 2014. Further Developments in the Right to be Forgotten: The European Court of Justice's Judgment in Case C-131/12, Google Spain, SL, Google Inc v Agencia Espanola de Proteccion de Datos. *Hum. Rts. L. Rev.*, **14**, 761.

Frické, Martin. 2009. The knowledge pyramid: a critique of the DIKW hierarchy. *Journal of information science*, **35**(2), 131–142.

Fuchs, Christian. 2012. Google capitalism. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, **10**(1), 42–48.

Fuchs, Christian. 2013. *Social media: A critical introduction*. Sage.

Fuller, Matthew. 2003. *Behind the blip: Essays on the culture of software*. Autonomedia.

Galloway, Alexander R. 2004. *Protocol: How control exists after decentralization*. MIT press.

Gibson, James J. 2014. *The ecological approach to visual perception: classic edition*. Psychology Press.

Gillespie, Tarleton. 2010. The politics of 'platforms'. *New Media & Society*, **12**(3), 347–364.

Gillespie, Tarleton. 2014. The Relevance of Algorithms. *Media technologies: Essays on communication, materiality, and society*, 167–194.

Goffman, Erving. 1959. *The presentation of self in everyday life*. Garden City, NY Double Day.

Goffman, Erving. 1963. *Stigma: Notes on a spoiled identity*. Penguin Books.

Goldhaber, Michael H. 1997. The attention economy and the net. *First Monday*, **2**(4).

Goldman, Eric. 2011. Revisiting search engine bias. *William Mitchell Law Review*, **38**(1), 96–110.

Gorzeman, Ludo, & Korenhof, Paulan. 2016. Escaping the Panopticon Over Time. *Philosophy & Technology*, 1–20.

Graux, Hans, Ausloos, Jef, & Valcke, Peggy. 2012. The right to be Forgotten in the Internet Era. *ICRI research paper*.

Gregory, Sam, & Losh, Elizabeth. 2012. Remixing human rights: Rethinking civic expression, representation and personal security in online video. *First Monday*, **17**(8).

Grimmelmann, James. 2010a. The privacy virus. *Chap. 1, pages 3–12 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Grimmelmann, James. 2010b. Some skepticism about search neutrality. *The next digital decade: Essays on the future of the Internet*, 435–459.

Gross, Ralph, & Acquisti, Alessandro. 2005. Information revelation and privacy in online social networks. *Pages 71–80 of: Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. ACM.

Gulli, Antonio, & Signorini, Alessio. 2005. The indexable web is more than 11.5 billion pages. *Pages 902–903 of: Special interest tracks and posters of the 14th international conference on World Wide Web*. ACM.

Haimson, Oliver, & Hoffmann, Anna. 2016. Constructing and enforcing äuthenticïdentity online: Facebook, real names, and non-normative identities. *First Monday*, **21**(6).

Hamington, Maurice. 2010. Care Ethics, Friendship, and Facebook. *Chap. 12, pages 135–145 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Haraway, Donna. 1991. *Simians, Cyborgs, and Women*. Routledge.

Hargittai, Eszter. 2000. Open portals or closed gates? Channeling content on the World Wide Web. *Poetics*, **27**(4), 233–253.

Hargittai, Eszter. 2003. Informed web surfing: The social context of user sophistication. *Society online: The Internet in context*, 257–74.

Harjumaa, Marja, & Oinas-Kukkonen, Harri. 2007. Persuasion theories and IT design. *Pages 311–314 of: International Conference on Persuasive Technology*. Springer.

Heath, Tom, & Motta, Enrico. 2008. Ease of interaction plus ease of integration: Combining Web2. 0 and the Semantic Web in a reviewing site. *Web Semantics: Science, Services and Agents on the World Wide Web*, **6**(1), 76–83.

Heersmink, Richard. 2012. Defending extension theory: A response to Kiran and Verbeek. *Philosophy & Technology*, **25**(1), 121–128.

Heidegger, Martin. 1977. *The question concerning technology and other essays.* New York: Harper & Row.

Helberger, Natali, & van Hoboken, Joris. 2010. Little Brother is Tagging You-Legal and Policy Implications of Amateur Data Controllers. *Computer Law International (CRi)*, 101–109.

Hendler, James, Shadbolt, Nigel, Hall, Wendy, Berners-Lee, Tim, & Weitzner, Daniel. 2008. Web science: an interdisciplinary approach to understanding the web. *Communications of the ACM*, **51**(7), 60–69.

Hess, Aaron. 2008. Reconsidering the rhizome: A textual analysis of web search engines as gatekeepers of the internet. *Pages 35–50 of: Web search*. Springer.

Hijmans, Hielke. 2014. Right to Have Links Removed: Evidence of Effective Data Protection: Case C-131/12 Google v. Agencia Española de Protectión de Datos (AEPD) and Mario Costeja Gonzalez, Judgment of 13 May 2014. *Maastricht Journal of European and Comparative Law*, **21**(3), 555–563.

Hildebrandt, M. 2009. Where idem meets ipse: Conceptual analysis. *Where idem-identity meets ipse-identity. Conceptual explorations*, 12–17.

Hildebrandt, Mireille. 2006. Privacy and identity. *Privacy and the criminal law*, **43**.

Hildebrandt, Mireille. 2015. *Smart Technologies and the End (s) of Law: Novel Entanglements of Law and Technology*. Edward Elgar Publishing.

Hill, David W. 2009. Reflections on leaving Facebook. *Fast Capitalism*, **5**(2).

Hill, Robin. 2016. What an Algorithm Is. *Philosophy and Technology*, **29:35-59**, 25.

Hinman, Lawrence M. 2008. Searching ethics: The role of search engines in the construction and distribution of knowledge. *Pages 67–76 of: Web search*. Springer.

Hoffman, David, Bruening, Paula, & Carter, Sophia. 2015. The right to obscurity: How we can implement the google spain decision. *NCJL & Tech.*, **17**, 437.

Holman, E Alison, & Silver, Roxane Cohen. 1998. Getting "stuck" in the past: temporal orientation and coping with trauma. *Journal of personality and social psychology*, **74**(5), 1146.

Hoskins, Andrew. 2014. The right to be forgotten in post-scarcity culture. *Pages 50–64 of: The Ethics of Memory in a Digital Age*. Springer.

Hui, Yuk. 2013. What is a digital object? *Chap. 4, pages 52–67 of:* Halpin, Harry, & Monnin, Alexandre (eds), *Philosophical Engineering: Toward a Philosophy of the Web*. John Wiley & Sons.

Hui, Yuk. 2016. On the Synthesis of Social Memories. *Pages 307–325 of: Memory in Motion: Archives, Technology and the Social*. Amsterdam University Press.

Husserl, Edmund. 1991. On the Phenomenology of the Consciousness of Internal Time, trans. by JB Brough. *Kluwer Academic*, **39**, 84–88.

Ibanez, Maria-Blanca, Di-Serio, Angela, & Delgado-Kloos, Carlos. 2014. Gamification for engaging computer science students in learning activities: A case study. *IEEE Transactions on learning technologies*, **7**(3), 291–301.

Iedema, Rick. 2001. Resemiotization. *Semiotica*, **2001**(137).

Iglezakis, Ioannis. 2016. The Right To Be Forgotten: A New Digital Right for cyberspace. *EDIÇÃO N. º III–FEVEREIRO DE 2017*, **17**, 67.

Ihde, Don. 1983. *Existential technics*. SUNY Press.

Ihde, Don. 1990. *Technology and the lifeworld: From garden to earth*. Indiana University Press.

Ippolita. 2015. *In the Facebook Aquarium: The Resistible Rise of Anarcho-capitalism*. Institute of Network Cultures.

Jamieson, Jack. 2016. Many (to platform) to many: Web 2.0 application infrastructures. *First Monday*, **21**(6).

Jappy, Tony. 2013. *Introduction to Peircean Visual Semiotics*. A&C Black.

Jiang, Lu, Miao, Yajie, Yang, Yi, Lan, Zhenzhong, & Hauptmann, Alexander G. 2014. Viral video style: A closer look at viral videos on youtube. *Page 193 of: Proceedings of International Conference on Multimedia Retrieval*. ACM.

John, Nancy R. 1996. Putting content onto the Internet. *First Monday*, **1**(2).

Jones, Meg Leta. 2018. *Ctrl+ Z: The right to be forgotten*. NYU Press.

Jones, Meg Leta, Zeide, Elana, Mai, Jens-Erik, Jones, Elisabeth, Dupre, Jill, & Richards, Neil. 2015. The right to be forgotten. *Proceedings of the Association for Information Science and Technology*, **52**(1), 1–3.

Kampmark, Binoy. 2015. To find or be forgotten: Global tensions on the right to erasure and internet governance. *Journal of Global Faultlines*, **2**(2), 1–18.

Kaplan, Frederic. 2014. Linguistic capitalism and algorithmic mediation. *Representations*, **127**(1), 57–63.

Karapapa, Stavroula, & Borghi, Maurizio. 2015. Search engine liability for autocomplete suggestions: personality, privacy and the power of the algorithm. *International Journal of Law and Information Technology*, **23**(3), 261–289.

Kelly, Kevin. 2007. Scan this book! *Chap. Scan This Book!, pages 69–93 of:* Levy, Steven (ed), *The Best of Technology Writing*. The University of Michigan Press.

Kiel, Joan M. 2005. The digital divide: Internet and e-mail use by the elderly. *Medical Informatics and the Internet in Medicine*, **30**(1), 19–23.

Kim, Yoojung, Sohn, Dongyoung, & Choi, Sejung Marina. 2011. Cultural difference in motivations for using social network sites: A comparative study of American and Korean college students. *Computers in human behavior*, **27**(1), 365–372.

King, Nicola. 2000. *Memory, Narrative, Identity: Remembering the Self*. Edinburgh University Press.

Kinsley, Samuel. 2015. Memory programmes: the industrial retention of collective life. *cultural geographies*, **22**(1), 155–175.

Kiran, Asle H, & Verbeek, Peter-Paul. 2010. Trusting our selves to technology. *Knowledge, Technology & Policy*, **23**(3-4), 409–427.

Kitchin, Rob. 2017. Thinking critically about and researching algorithms. *Information, Communication & Society*, **20**(1), 14–29.

Knight, SA, & Spink, Amanda. 2008. Toward a web search information behavior model. *Pages 209–234 of: Web search*. Springer.

König, René, & Rasch, Miriam. 2014. Reflect and Act! Introduction to the Society of the Query Reader. *Society of the Query Reader. Reflections on Web Search, Amsterdam, Institute of Network Cultures*, 9–16.

Koops, Bert-Jaap. 2011. Forgetting footprints, shunning shadows: A critical analysis of the right to be forgotten in big data practice. *SCRIPTed*, **8**, 229.

Koops, Bert-Jaap. 2014. The trouble with European data protection law. *International Data Privacy Law*, **4**(4), 250–261.

Koops, Bert-Jaap, Newell, Bryce Clayton, Timan, Tjerk, Skorvanek, Ivan, Chokrevski, Tomislav, & Galic, Masa. 2017. A typology of privacy. *University of Pennsylvania Journal of International Law*, **38**, 483–575.

Korenhof, Paulan. 2013. Forgetting Bits and Pieces: An Exploration of the 'Right to Be Forgotten' as Implementation of 'Forgetting' in Online Memory Processes. *Pages 114–127 of: IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*. Springer.

Korenhof, Paulan. 2014. Stage ahoy! Deconstruction of the "Drunken Pirate" Case in the Light of Impression Management. *Pages 79–97 of: Reloading Data Protection*. Springer.

Korenhof, Paulan, & Koops, Bert-Jaap. 2013. Gender Identity and Privacy: Could a Right to Be Forgotten Help Andrew Agnes Online? *Chap. 6 of:* Ghezzi, Alessia, Guimarães Pereira, Ângela, & Vesnić-Alujević, Lucia (eds), *The ethics of memory in a digital age; interrogating the right to be forgotten*. Palgrave Macmillan.

Korenhof, Paulan, Ausloos, Jef, Szekely, Ivan, Ambrose, Meg, Sartor, Giovanni, & Leenes, Ronald. 2015. Timing the right to be forgotten: A study into "time" as a factor in deciding about retention or erasure of data. *Pages 171–201 of: Reforming European data protection law*. Springer.

Koskela, Hille. 2004. Webcams, TV shows and mobile phones: Empowering exhibitionism. *Surveillance & Society*, **2**(2/3).

Kuipers, Giselinde. 2009. Humor styles and symbolic boundaries. *Journal of Literary Theory*, **3**(2), 219–239.

Kulk, Stefan, & Borgesius, Frederik Zuiderveen. 2014. Google Spain v. González: Did the Court Forget about Freedom of Expression?: Case C-131/12 Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos and Mario Costeja González. *European Journal of Risk Regulation*, **5**(3), 389–398.

Kulk, Stefan, & Borgesius, Frederik Zuiderveen. 2018. Privacy, freedom of expression, and the right to be forgotten in Europe. *Cambridge Handbook of Consumer Privacy*, 301.

Kulk, Stefan, & Zuiderveen Borgesius, Frederik. 2015. De implicaties van het Google Spain-arrest voor de vrijheid van meningsuiting (The Implications of the Google Spain Judgment for Freedom of Expression). *NTM/NJCM-Bulletin (2015)*, **40**(1), 3–19.

Kulk, Stefan, & Zuiderveen Borgesius, Frederik. 2017. Annotatie bij Hof Den haag 23 mei 2017, nr. 200.194.334/01. *Computerrecht*, **204**(5), 309–318.

Kuner, Christopher. 2014. The Court of Justice of the EU judgment on data protection and internet search engines: Current Issues and Future Challenges. *Pages 19–55 of:* Hess, Burkhard, & Mariottini, Cristina M. (eds), *Protecting Privacy in Private International and Procedural Law and by Data Protection*.

Kuner, Christopher. 2015. Extraterritoriality and regulation of international data transfers in EU data protection law. *International Data Privacy Law*, **5**(4), 235–245.

Kupfer, Joseph. 1987. Privacy, autonomy, and self-concept. *American Philosophical Quarterly*, **24**(1), 81–89.

Langlois, Ganaele. 2013. Participatory culture and the new governance of communication: The paradox of participatory media. *Television & New Media*, **14**(2), 91–105.

Larson III, Robert G. 2013. Forgetting the First Amendment: How obscurity-based privacy and a right to be forgotten are incompatible with free speech. *Communication Law and Policy*, **18**(1), 91–120.

Latour, Bruno. 1993. *We have never been modern*. Harvard University Press.

Lazzarato, Maurizio. 2014. *Signs and machines: Capitalism and the production of subjectivity*. Semiotext(e) Los Angeles.

Lee, Eunji, Lee, Jung-Ah, Moon, Jang Ho, & Sung, Yongjun. 2015. Pictures speak louder than words: Motivations for using Instagram. *Cyberpsychology, Behavior, and Social Networking*, **18**(9), 552–556.

Lee, Eunsun, Ahn, Jungsun, & Kim, Yeo Jung. 2014. Personality traits and self-presentation at Facebook. *Personality and Individual Differences*, **69**, 162–167.

Lee, Zuk-Nae. 1999. Korean culture and sense of shame. *Transcultural psychiatry*, **36**(2), 181–194.

Lee-Won, Roselyn J, Shim, Minsun, Joo, Yeon Kyoung, & Park, Sung Gwan. 2014. Who puts the best "face" forward on Facebook?: Positive self-presentation in online social networking and the role of self-consciousness, actual-to-total Friends ratio, and culture. *Computers in Human Behavior*, **39**, 413–423.

Leenes, Ronald. 2009. Context is everything sociality and privacy in online social network sites. *Pages 48–65 of: IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*. Springer.

Leenes, Ronald. 2011. Framing techno-regulation: An exploration of state and non-state regulation by technology. *Legisprudence*, **5**(2), 143–169.

Lemmens, Pieter. 2014. Social Autonomy and Heteronomy in the Age of ICT: The Digital Pharmakon and the (Dis) Empowerment of the General Intellect. *Foundations of Science*, 1–10.

Lemmens, Pieter. 2015. Van Ontologie Naar Organologie: Heidegger En Stiegler over Het Gevaar En De Ambivalentie Van De Techniek. *De Uil van Minerva*, **28**(4), 335–362.

Leone, Massimo. 2018. Semiotics of the Selfie: Glorification of the Present. *Punctum*, 33–48.

Leroi-Gourhan, André. 1993. *Gesture and speech*. MIT Press.

Lessig, Lawrence. 2006. *Code Version 2.0*. Basic Books.

Li, Wenlong. 2018. A tale of two rights: exploring the potential conflict between right to data portability and right to be forgotten under the General Data Protection Regulation. *International Data Privacy Law*, **8**(4), 309–317.

Lindsay, David. 2014. The 'right to be forgotten'by search engines under data privacy law: A legal analysis of the Costeja ruling. *Journal of Media Law*, **6**(2), 159–179.

Losee, Robert M. 1997. A discipline independent definition of information. *Journal of the American Society for Information Science (1986-1998)*, **48**(3), 254–269.

Losh, Elizabeth. 2010. With Friends Like These, Who Needs Enemies? *Chap. 4, pages 33–47 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Lovink, Geert. 2005. *The Principle of Notworking, Concepts in Critical Internet Culture*. HvA Publicaties.

Lovink, Geert. 2016. *Social Media Abyss: Critical Internet Cultures and the Force of Negation*. John Wiley & Sons.

Low, David William. 2009. Data, information, knowledge: a semiotic-system's view for database design. *In: Proceedings of the 53rd Annual Meeting of the ISSS-2009, Brisbane, Australia*, vol. 1.

Lyndon, Amy, Bonds-Raacke, Jennifer, & Cratty, Alyssa D. 2011. College students' Facebook stalking of ex-partners. *Cyberpsychology, Behavior, and Social Networking*, **14**(12), 711–716.

Lynskey, Orla. 2013. Time to forget the 'Right to be Forgotten'? Advocate General Jääskinen's opinion in C-131/12 Google Spain v AEPD. *European Law Blog*, **3**.

Lynskey, Orla. 2015. Control over Personal Data in a Digital Age: Google S pain v AEPD and Mario Costeja Gonzalez. *The Modern Law Review*, **78**(3), 522–534.

Ma, Xiao, Hancock, Jeff, & Naaman, Mor. 2016. Anonymity, intimacy and self-disclosure in social media. *Pages 3857–3869 of: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM.

Mager, Astrid. 2012. Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication & Society*, **15**(5), 769–787.

Manoff, Marlene. 2010. Archive and database as metaphor: Theorizing the Historical Record. *portal: Libraries and the Academy*, **10**(4), 385–398.

Manovich, Lev. 2001. *The language of new media*. MIT press.

Mantelero, Alessandro. 2013. The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'. *Computer Law & Security Review*, **29**(3), 229–235.

Marcuse, Herbert. 1966. *One Dimensional Man: Studies in Idiology of Avanced Industrial Society*. Beacon Press Boston.

Markou, Christiana. 2015. The 'Right to Be Forgotten': Ten reasons why it should be forgotten. *Pages 203–226 of: Reforming European Data Protection Law*. Springer.

Mayer, Katja. 2009. On the Sociometry of Search Engines. *Deep Search. The Politics of Search Beyond Google*, 54–72.

Mayer-Schönberger, Viktor. 2009. *Delete: The virtue of forgetting in the digital age*. Princeton University Press.

McAdams, Dan P, Josselson, Ruthellen Ed, & Lieblich, Amia Ed. 2006. *Identity and story: Creating self in narrative*. American Psychological Association.

McDonald, Steven. 2019. The Right to be Forgotten: The Potential Effects on Canadian. *Dalhousie Journal of Interdisciplinary Management*, **15**.

McGrenere, Joanna, & Ho, Wayne. 2000. Affordances: Clarifying and evolving a concept. *Pages 179–186 of: Graphics interface*, vol. 2000.

Meikle, Graham. 2010. It's like talking to a wall. *Chap. 2, pages 13–20 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Merton, Robert K, *et al.* 1968. The Matthew effect in science. *Science*, **159**(3810), 56–63.

Miconi, Andrea. 2014. Dialectic of Google. *Society of the Query Reader. Reflections on Web Search, Amsterdam, Institute of Network Cultures*, 30–40.

Middleton, David, & Brown, Steve D. 2005. *The social psychology of experience: Studies in remembering and forgetting*. Sage.

Miltner, Kate M. 2014. "There's no place for lulz on LOLCats": The role of genre, gender, and group identity in the interpretation and enjoyment of an Internet meme. *First Monday*, **19**(8).

Mink, Louis O. 1978. Narrative form as a cognitive instrument. *The writing of history: Literary form and historical understanding*, 129–149.

Mitchell, Liam. 2014. Life on automatic: Facebook's archival subject. *First Monday*, **19**(2).

Mitrou, Lilian, & Karyda, Maria. 2012. *EU's data protection reform and the right to be forgotten: A legal response to a technological challenge?* `http://www.icsd.aegean.gr/website_files/proptyxiako/388450775.pdf`. Last accessed: 01-11-2018.

Mittelstadt, Brent Daniel, Allo, Patrick, Taddeo, Mariarosaria, Wachter, Sandra, & Floridi, Luciano. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society*, **3**(2), 1–12.

Moerel, L, & Prins, C. 2015. On the Death of Purpose Limitation. *Privacy Perspectives Where the Real Conversation in Privacy Happens*, **2**.

Moore, Gordon E. 2006. Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp. 114 ff. *IEEE Solid-State Circuits Society Newsletter*, **11**(3), 33–35.

Morrison, Stacey, & Gomez, Ricardo. 2014. Pushback: the growth of expressions of resistance to constant online connectivity. *iConference 2014 Proceedings*.

Moulier-Boutang, Yann. 2011. *Cognitive capitalism*. Polity Press.

Nabi, Zubair. 2014. Censorship is futile. *First Monday*, **19**(11).

Nagy, Peter, & Neff, Gina. 2015. Imagined Affordance: Reconstructing a Keyword for Communication Theory. *Social Media+ Society*, **1**(2), 1–9.

Nahon, Karine, & Hemsley, Jeff. 2013. *Going viral*. Polity Press.

Negri, Antonio. 2005. *The politics of subversion: A manifesto for the twenty-first century*. Polity Press.

Nieuwenhuis, Aernout J. 2011. Over de grens van de vrijheid van meningsuiting. *Nijmegen: Ars Aequi*.

Nissenbaum, Helen. 2004. Privacy as contextual integrity. *Wash. L. Rev.*, **79**, 119.

Nissenbaum, Helen. 2010. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Nissenbaum, Helen, & Brunton, Finn. 2015. *Obfuscation: A user's guide for privacy and protest*. Cambridge, MA: MIT Press.

Nöth, Winfried. 2011. Representation and reference according to Peirce. *International Journal of Signs and Semiotic Systems (IJSSS)*, **1**(2), 28–39.

Novotny, Alexander, & Spiekermann, Sarah. 2014. *Oblivion on the web: an inquiry of user needs and technologies.* `https://epub.wu.ac.at/4112/1/NovotnySpiekermann2014_OblivionOnTheWeb.pdf`. Last accessed: 31-10-2019.

Obar, Jonathan A, & Oeldorf-Hirsch, Anne. 2018. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 1–20.

Obendorf, Hartmut, & Weinreich, Harald. 2003. Comparing link marker visualization techniques: changes in reading behavior. *Pages 736–745 of: Proceedings of the 12th international conference on World Wide Web*. ACM.

O'Callaghan, Patrick, & de Mars, Sylvia. 2016. Narratives about privacy and forgetting in English law. *International Review of Law, Computers & Technology*, **30**(1-2), 42–56.

O'Hara, Kieron. 2015. The right to be forgotten: the good, the bad, and the ugly. *IEEE Internet Computing*, **19**(4), 73–79.

Oinas-Kukkonen, Harri, & Harjumaa, Marja. 2008. Towards deeper understanding of persuasion in software and information systems. *Pages 200–205 of: First international conference on advances in computer-human interaction*. IEEE.

Oravec, Jo Ann. 2002. Bookmarking the world: Weblog applications in education. *Journal of Adolescent & Adult Literacy*, **45**(7), 616–621.

Origgi, Gloria. 2012. Designing wisdom through the web. Reputation and the passion of ranking. *Collective Wisdom: Principles and Mechanisms*, 38–55.

Oudshoorn, Nelly, & Pinch, Trevor. 2003. How users and non-users matter. *Chap. Introduction, pages 1–25 of:* Oudshoorn, Nelly, & Pinch, Trevor (eds), *How users matter: the co-construction of users and technology (inside technology)*. MIT Press.

Oudshoorn, Nelly, Saetnan, Ann Rudinow, & Lie, Merete. 2002. On gender and things: Reflections on an exhibition on gendered artifacts. *Pages 471–483 of: Women's Studies International Forum*, vol. 25. Elsevier.

Oudshoorn, Nelly, Rommes, Els, & Stienstra, Marcelle. 2004. Configuring the user as everybody: Gender and design cultures in information and communication technologies. *Science, Technology & Human Values*, **29**(1), 30–63.

Oulasvirta, Antti, Rattenbury, Tye, Ma, Lingyi, & Raita, Eeva. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, **16**(1), 105–114.

Padova, Yann. 2019. Is the right to be forgotten a universal, regional, or 'glocal' right? *International Data Privacy Law*.

Page, Lawrence, Brin, Sergey, Motwani, Rajeev, & Winograd, Terry. 1999. The PageRank citation ranking: bringing order to the Web. *Stanford InfoLab*.

Pariser, Eli. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Park, Namsu, Kee, Kerk F, & Valenzuela, Sebastián. 2009. Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *CyberPsychology & Behavior*, **12**(6), 729–733.

Pasquale, Frank. 2015. *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Pasquinelli, Matteo. 2009. Google's PageRank algorithm: A diagram of cognitive capitalism and the rentier of the common intellect. *Deep search*, **3**, 152–162.

Peguera, Miquel. 2015. The shaky ground of the right to be delisted. *Vand. J. Ent. & Tech. L.*, **18**, 507.

Peirce, Charles Sanders. 1974. *Collected papers of charles sanders peirce*. Vol. 5. Harvard University Press.

Peirce, Charles Sanders. 1998. *The essential Peirce: selected philosophical writings*. Vol. 2. Indiana University Press.

Pellauer, David, & Dauenhauer, Bernard. 2016. Paul Ricoeur. *In:* Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, winter 2016 edn. Metaphysics Research Lab, Stanford University.

Perelman, Ch, & Olbrechts-Tyteca, L. 1969. *The new rhetoric. A treatise on argumentation (Translation of La nouvelle rhetorique. Traite de l'argumentation). Paris: Presses Universitaires de France (1958). Notre Dame.*

Pfeil, Ulrike, Arjan, Raj, & Zaphiris, Panayiotis. 2009. Age differences in online social networking–A study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior*, **25**(3), 643–654.

Plato. 1973. *Phaedrus and Letters VII and VIII. Translated with introdusctions by Walter Hamilton*. Penguin Books.

Politou, Eugenia, Michota, Alexandra, Alepis, Efthimios, Pocs, Matthias, & Patsakis, Constantinos. 2018a. Backups and the right to be forgotten in the GDPR: An uneasy relationship. *Computer Law & Security Review*, **34**(6), 1247–1257.

Politou, Eugenia, Alepis, Efthimios, & Patsakis, Constantinos. 2018b. Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity*.

Porcedda, Maria Grazia. 2017. The Recrudescence of 'Security v. Privacy'after the 2015 Terrorist Attacks, and the Value of 'Privacy Rights' in the European Union. *Pages 137–180 of: Rethinking Surveillance and Control*. Nomos Verlagsgesellschaft mbH & Co. KG.

Porcedda, Maria Grazia. 2018. On Boundaries. Finding the Essence of the Right to the Protection of Personal Data. *Chap. 12 of:* Leenes, Ronald, van Brakel, Rosamunde, Gutwirth, Serge, & De Hert, Paul (eds), *Data protection and privacy: The internet of bodies*. Bloomsbury Publishing.

Post, Robert C. 2017. Data privacy and dignitary privacy: Google Spain, the right to be forgotten, and the construction of the public sphere. *Duke LJ*, **67**, 981.

Powles, Julia, & Floridi, Luciano. 2014. A Manifesto for the Future of the 'Right to be Forgotten' Debate. *The Guardian*, **22-07-2014**.

Quah, Danny. 2003. Digital goods and the new economy. *CEPR Discussion Paper*.

Quan-Haase, Anabel, & Wellman, Barry. 2005. Local virtuality in an organization: Implications for community of practice. *Pages 215–238 of: Communities and technologies 2005*. Springer.

Quan-Haase, Anabel, & Young, Alyson L. 2010. Uses and gratifications of social media: A comparison of Facebook and instant messaging. *Bulletin of Science, Technology & Society*, **30**(5), 350–361.

Raffl, Celina, Hofkirchner, Wolfgang, Fuchs, Christian, & Schafranek, Matthias. 2011. *The Web as a Techno-Social System: The Emergence of Web 3.0*. na.

Ranquet, Marie. 2019. The right to be forgotten: A new fundamental right for the individual? *Communications*, 149–159.

Rao, Mithun Bantwal, Jongerden, Joost, Lemmens, Pieter, & Ruivenkamp, Guido. 2015. Technological mediation and power: Postphenomenology, critical theory, and autonomist marxism. *Philosophy & Technology*, **28**(3), 449–474.

Reich, Stephanie M, Subrahmanyam, Kaveri, & Espinoza, Guadalupe. 2012. Friending, IMing, and hanging out face-to-face: Overlap in adolescents' online and offline social networks. *Developmental psychology*, **48**(2), 356.

Reid, Philippa, Monsen, Jeremy, & Rivers, Ian. 2004. Psychology's contribution to understanding and managing bullying within schools. *Educational Psychology in Practice*, **20**(3), 241–258.

Ricoeur, Paul. 1976. *Interpretation theory: Discourse and the surplus of meaning*. TCU press.

Ricoeur, Paul. 1991. Narrative identity. *Philosophy today*, **35**(1), 73–81.

Ricoeur, Paul. 1992. *Oneself as another*. University of Chicago Press.

Ridenour, Sara. 2011. Facebook Killed the Reunion Star: How Facebook is Changing Who We Are and What We Do. *Fast Capitalism*, **8**(1).

Rodotà, Stefano. 2009. Data Protection as a Fundamental Right. *Pages 77–82 of:* Gutwirth, Serge, Poullet, Yves, De Hert, Paul, de Terwangne, Cécile, & Nouwt, Sjaak (eds), *Reinventing Data Protection?* Springer Netherlands.

Roessler, Beate. 2005. *The value of privacy*. Polity Press.

Ronson, Jon. 2016. *So you've been publicly shamed*. Riverhead Books.

Roosendaal, Arnold. 2009. Digital personae and profiles as representations of individuals. *Pages 226–236 of: IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*. Springer.

Roosendaal, Arnold. 2010. Facebook tracks and traces everyone: Like this! *Tilburg Law School Legal Studies Research Paper Series*, **2011**(03). Last accessed: 13-10-2019.

Rosen, Jeffrey. 2011. The right to be forgotten. *Stan. L. Rev. Online*, **64**, 88.

Rosen, Voir Jeffrey. 2010. The web means the end of forgetting. *The New York Times*, **21**.

Rosenberger, Robert, & Verbeek, Peter-Paul. 2015. A field guide to postphenomenology. *Postphenomenological investigations: Essays on human-technology relations*, 9–42.

Ross, Anthony. 2013. Distance and Presence in Analogue and Digital Epistolary Networks. *Techné: Research in Philosophy and Technology*.

Rouvroy, Antoinette, & Poullet, Yves. 2009. The right to informational self-determination and the value of self-development: Reassessing the importance of privacy for democracy. *Pages 45–76 of: Reinventing data protection?* Springer.

Rowan, Murray, Gregor, Peter, Sloan, David, & Booth, Paul. 2000. Evaluating web resources for disability access. *Pages 80–84 of: Proceedings of the fourth international ACM conference on Assistive technologies*. ACM.

Rowley, Jennifer E. 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*.

Rushkoff, Douglas. 2010. *Program or be programmed: Ten commands for a digital age*. Or Books.

Rustad, Michael L, & Kulevska, Sanna. 2014. Reconceptualizing the right to be forgotten to enable transatlantic data flow. *Harv. JL & Tech.*, **28**, 349.

Sanders, John T. 1993. Merleau-Ponty, Gibson, and the materiality of meaning. *Man and World*, **26**(3), 287–302.

Sarachan, Jeremy. 2010. Profile picture, right here, right now. *Chap. 5, pages 51–64 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Sarkar, Subhadeep, Banatre, Jean-Pierre, Rilling, Louis, & Morin, Christine. 2018. Towards Enforcement of the EU GDPR: Enabling Data Erasure. *Pages 222–229 of: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. IEEE.

Sartor, Giovanni. 2014. Search Engines as Controllers: Inconvenient Implications of a Questionable Classification: Case C-131/12 Google v. Agencia Española de Protectión de Datos (AEPD) and Mario Costeja Gonzalez, Judgment of 13 May 2014. *Maastricht Journal of European and Comparative Law*, **21**(3), 564–575.

Sartor, Giovanni. 2015. The right to be forgotten: balancing interests in the flux of time. *International journal of law and information technology*, **24**(1), 72–98.

Sartre, Jean-Paul. 1987. *Huis clos*. Psychology Press.

Schäfer, Mirko Tobias. 2011. *Bastard Culture! User participation and the extension of the cultural industries*. Amsterdam University Press.

Schoeman, Ferdinand David. 1984. *Philosophical dimensions of privacy: An anthology*. Cambridge University Press.

Seidman, Gwendolyn. 2013. Self-presentation and belonging on Facebook: How personality influences social media use and motivations. *Personality and Individual Differences*, **54**(3), 402–407.

Sen, Amartya. 2007. *Identity and violence: The illusion of destiny*. Penguin Books India.

Shannon, Claude. 1993. *Collected papers*. New York: IEEE Press.

Sharma, Nita A, & Kurhekar, Premlata P. 2013. Content Management System. *International Journal of Innovative Research and Development—— ISSN 2278–0211*, **2**(12).

Shifman, Limor. 2012. An anatomy of a YouTube meme. *New Media & Society*, **14**(2), 187–203.

Shifman, Limor. 2013. *Memes in digital culture*. MIT Press.

Short, Thomas Lloyd. 2007. *Peirce's theory of signs*. Cambridge University Press.

Silverstone, Roger, & Haddon, Leslie. 1996. *Design and the domestication of information and communication technologies: Technical change and everyday life*. Oxford University Press.

Simon, Herbert A. 1969. Designing organizations for an information-rich world. *Chap. 2, pages 37–72 of:* Greenberger, M. (ed), *Computers, Communications, and the Public Interest*. The Johns Hopkins Press.

Singer, Jane B. 2009. Convergence and divergence. *Journalism*, **10**(3), 375–377.

Singleton, Shaniqua. 2015. Balancing a Right to be Forgotten with a Right to Freedom of Expression in the Wake of Google Spain v. AEPD. *Ga. J. Int'l & Comp. L.*, **44**, 165.

Solove, Daniel J. 2005. A taxonomy of privacy. *University of Pennsylvania Law Review*, **154**, 477.

Solove, Daniel J. 2007. *The future of reputation: Gossip, rumor, and privacy on the Internet*. Yale University Press.

Sorokowska, Agnieszka, Oleszkiewicz, Anna, Frackowiak, Tomasz, Pisanski, Katarzyna, Chmiel, Anna, & Sorokowski, Piotr. 2016. Selfies and personality: Who posts self-portrait photographs? *Personality and Individual Differences*, **90**, 119–123.

Sparrow, B Chatman, Chapman, P, & Gould, J. 2005. L.(2013). *Social cognition in the internet age: Same as it ever was?*, 273–292.

Sparrow, Betsy, Liu, Jenny, & Wegner, Daniel M. 2011. Google effects on memory: Cognitive consequences of having information at our fingertips. *science*, **333**(6043), 776–778.

Stalder, Felix, & Mayer, Christine. 2009. The Second Index Search Engines, Personalization and Surveillance. *Pages 98–115 of:* Becker, Konrad, & Stalder, Felix (eds), *Deep Search. The Politics of Search beyond Google*. Studienverlag.

Stallybrass, Peter. 2007. Against thinking. *PMLA*, **122**(5), 1580–1587.

Stanfill, Mel. 2015. The interface as discourse: The production of norms through web design. *new media & society*, **17**(7), 1059–1074.

Stiegler, Bernard. 1998. *Technics and time: The fault of Epimetheus*. Vol. 1. Stanford University Press.

Stiegler, Bernard. 2009. *Technics and Time: Disorientation*. Vol. 2. Stanford, California: Stanford University Press.

Stiegler, Bernard. 2010a. Anamnesis and hypomnesis. *Ars industrialis*, **20**.

Stiegler, Bernard. 2010b. *For a new critique of political economy*. Polity Press.

Stiegler, Bernard. 2011. *Decadence of industrial democracies*. Vol. 1. Polity Press.

Stiegler, Bernard. 2012. Relational ecology and the digital pharmakon. *Culture Machine*, **13**.

Stiegler, Bernard. 2014. *Symbolic Misery: The Hyperindustrial Epoch*. Vol. 1. Polity Press.

Stuart, Allyson Haynes. 2013. Google search results: buried if not forgotten. *NCJL & Tech.*, **15**, 463.

Stute, David J. 2014. Privacy Almighty-The CJEU's Judgment in Google Spain SL v. AEPD. *Mich. J. Int'l L.*, **36**, 649.

Suler, John. 2004. The online disinhibition effect. *Cyberpsychology & behavior*, **7**(3), 321–326.

Sung, Yongjun, Lee, Jung-Ah, Kim, Eunice, & Choi, Sejung Marina. 2016. Why we post selfies: Understanding motivations for posting pictures of oneself. *Personality and Individual Differences*, **97**, 260–265.

Susser, Daniel. 2016. Information Privacy and Social Self-Authorship. *Techné: Research in Philosophy and Technology*, **20**(3), 216–239.

Svantesson, Dan Jerker B. 2015. Extraterritoriality and targeting in EU data privacy law: the weak spot undermining the regulation. *International Data Privacy Law*, **5**(4), 226–234.

Sweeney, Latanya. 2000. Simple demographics often identify people uniquely. *Health (San Francisco)*, **671**, 1–34.

Sweeney, Latanya. 2013. Discrimination in online ad delivery. *Queue*, **11**(3), 10.

Szekely, Ivan. 2012. The right to forget, the right to be forgotten. *Pages 347–363 of: European Data Protection: In Good Health?* Springer.

Szekely, Ivan. 2014. The right to be forgotten and the new archival paradigm. *Pages 28–49 of: The Ethics of Memory in a Digital Age*. Springer.

Taddicken, Monika. 2014. The 'Privacy Paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, **19**(2), 248–273.

Tavani, Herman. 2016. Search Engines and Ethics. *In:* Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, fall 2016 edn. Metaphysics Research Lab, Stanford University.

Tavani, Herman T. 2018. Should We Have a Right to Be Forgotten?: A Critique of Key Arguments Underlying This Question. *Journal of Information Ethics*, **27**(2), 26.

Teixeira, Thales. 2012. The new science of viral ads. *Harvard Business Review*, **90**(3 (March 2012)), 25–27.

Terranova, Tiziana. 2012. Attention, economy and the brain. *Culture Machine*, **13**(1), 1–19.

Thaler, Richard H, & Sunstein, Cass R. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thalos, Mariam. 2010. Why I am not a Friend. *Chap. 7, pages 75–88 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Thompson, John B. 2005. The new visibility. *Theory, Culture & Society*, **22**(6), 31–51.

Timan, Tjerk, & Oudshoorn, Nelly. 2012. Mobile cameras as new technologies of surveillance? How citizens experience the use of mobile cameras in public nightscapes. *Surveillance & Society*, **10**(2), 167.

Tirosh, Noam. 2017. Reconsidering the 'Right to be Forgotten'–memory rights and the right to memory in the new media era. *Media, Culture & Society*, **39**(5), 644–660.

Toma, Catalina L, & Hancock, Jeffrey T. 2013. Self-affirmation underlies Facebook use. *Personality and Social Psychology Bulletin*, **39**(3), 321–331.

Treadaway, Chris, & Smith, Mari. 2012. *Facebook marketing: An hour a day*. John Wiley & Sons.

Tredinnick, Luke. 2008. *Digital information culture: the individual and society in the digital age*. Elsevier.

Trottier, Daniel. 2011. A research agenda for social media surveillance. *Fast Capitalism*, **8**(1).

Tuomi, Ilkka. 1999. Data is more than knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory. *Pages 12–pp of: Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on*. IEEE.

Vafopoulos, Michalis. 2013. Being, space, and time on the Web. *Chap. 6, pages 77–96 of:* Halpin, Harry, & Monnin, Alexandre (eds), *Philosophical Engineering: Toward a Philosophy of the Web*. John Wiley & Sons.

Vaidhyanathan, Siva. 2008. Naked in the Nonopticon. *The Chronicle of Higher Education*, **54**(23), B57.

Vaidhyanathan, Siva. 2012. *The Googlization of everything:(and why we should worry)*. Univ of California Press.

van Alsenoy, Brendan. 2017. Reconciling the (extra) territorial reach of the GDPR with public international law. *In:* Vermeulen, Gert, & Lievens, Eva (eds), *Data Protection and Privacy under Pressure. Transatlantic tensions, EU surveillance, and big data*. Maklu.

van Couvering, Elizabeth. 2008. The history of the Internet search engine: Navigational media and the traffic commodity. *Pages 177–206 of: Web search*. Springer.

van den Berg, Bibi, & Leenes, Ronald. 2010. Audience segregation in social network sites. *Pages 1111–1116 of: Social Computing (SocialCom), 2010 IEEE Second International Conference on*. IEEE.

van den Berg, Bibi, & Leenes, Ronald. 2011. Keeping up appearances: Audience segregation in social network sites. *Pages 211–231 of: Computers, Privacy and Data Protection: an Element of Choice*. Springer.

van den Berg, Bibi, & Leenes, Ronald E. 2013. Abort, retry, fail: scoping techno-regulation and other techno-effects. *Pages 67–87 of: Human law and computer law: Comparative perspectives*. Springer.

van Deursen, Alexander JAM, & van Dijk, Jan AGM. 2009. Using the Internet: Skill related problems in users' online behavior. *Interacting with Computers*, **21**(5-6), 393–402.

van Dijck, José. 2013. *The culture of connectivity: A critical history of social media*. Oxford University Press.

van Hoboken, Joris. 2011. 9 Reasons Why a 'Right to Be Forgotten'Is Really Wrong. *Blog post, December*, **8**. Last accessed: 24-09-2019.

van Hoboken, Joris. 2012. *Search engine freedom: On the implications of the right to freedom of expression for the legal governance of web search engines*. Kluwer Law International Den Haag.

van Hoboken, Joris. 2013. The Proposed Right to Be Forgotten Seen from the Perspective of Our Right to Remember. *Freedom of Expression Safeguards in a Converging Information Environment, Prepared for the European Commission, Amsterdam*.

van House, Nancy A. 2011. Personal photography, digital technologies and the uses of the visual. *Visual Studies*, **26**(2), 125–134.

van Oost, Elizabeth CJ, Oudshoorn, NEJ, & Pinch, T. 2003. Materialized gender: how shavers configure the users' feminity and masculinity. *How users matter. The co-construction of users and technology*, 193–208.

Varis, Piia, & Blommaert, Jan. 2015. Conviviality and collectives on social media: Virality, memes, and new social structures. *Multilingual Margins: A journal of multilingualism from the periphery*, **2**(1), 31–31.

Vavra, Ashley Nicole. 2018. The right to be forgotten: An archival perspective. *The American Archivist*, **81**(1), 100–111.

Vejby, Rune, & Wittkower, Dylan. 2010. Spectacle 2.0. *Chap. 9, pages 97–108 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Verbeek, Peter-Paul. 2005. *What things do: Philosophical reflections on technology, agency, and design*. Penn State Press.

Verbeek, Peter-Paul. 2011. *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.

Virno, Paolo. 2003. *A Grammar of the Multitude*. Semiotext(e) Los Angeles.

Voigt, Paul, & Von dem Bussche, Axel. 2017. The EU General Data Protection Regulation (GDPR). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*.

Voss, W Gregory, & Castets-Renard, Céline. 2015. Proposal for an International Taxonomy on the Various Forms of the Right to Be Forgotten: A Study on the Convergence of Norms. *Colo. Tech. LJ*, **14**, 281.

Walter, Chip. 2005. Kryder's law. *Scientific American*, **293**(2), 32–33.

Wandel, Tamara, & Beavers, Anthony. 2010. Playing around with identity. *Chap. 8, pages 89–96 of:* Wittkower, D.E. (ed), *Facebook and Philosophy: What's on your Mind*. Chicago and La Salle: Open Court.

Wang, Dan, Xiang, Zheng, & Fesenmaier, Daniel R. 2016. Smartphone use in everyday life and travel. *Journal of Travel Research*, **55**(1), 52–63.

Wang, Jin-Liang, Jackson, Linda A, Wang, Hai-Zhen, & Gaskin, James. 2015. Predicting social networking site (SNS) use: Personality, attitudes, motivation and internet self-efficacy. *Personality and Individual Differences*, **80**, 119–124.

Wang, Yang, Norcie, Gregory, Komanduri, Saranga, Acquisti, Alessandro, Leon, Pedro Giovanni, & Cranor, Lorrie Faith. 2011. I regretted the minute I pressed share: A qualitative study of regrets on Facebook. *Page 10 of: Proceedings of the seventh symposium on usable privacy and security*. ACM.

Ward, David, Hahn, Jim, & Feist, Kirsten. 2012. Autocomplete as a research tool: a study on providing search suggestions. *Information Technology and Libraries (Online)*, **31**(4), 14.

Warren, Samuel D, & Brandeis, Louis D. 1890. The right to privacy. *Harvard law review*, 193–220.

Weber, Rolf H. 2011. The right to be forgotten: More than a Pandora's box. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, **2**, 120.

Wegner, Daniel M. 1987. Transactive memory: A contemporary analysis of the group mind. *Pages 185–208 of: Theories of group behavior*. Springer.

Wei, Carolyn Y, Evans, Mary B, Eliot, Matthew, Barrick, Jennifer, Maust, Brandon, & Spyridakis, Jan H. 2005. Influencing web-browsing behavior with intriguing and informative hyperlink wording. *Journal of information science*, **31**(5), 433–445.

Weinmann, Markus, Schneider, Christoph, & vom Brocke, Jan. 2016. Digital nudging. *Business & Information Systems Engineering*, **58**(6), 433–436.

Wellman, Barry. 2001. Little boxes, glocalization, and networked individualism. *Pages 10–25 of: Kyoto Workshop on Digital Cities*. Springer.

West, Tyler. 2011. Going viral: Factors that lead videos to become internet phenomena. *Elon Journal of Undergraduate Research*, **2**(1), 76–84.

Westin, Alan F. 1970. *Privacy and freedom*. Atheneum New York.

White, Peter, & White, Naomi. 2005. Virtually there: Travelling with new media. *First Monday*, **10**(8).

Wiener, Norbert. 1954. *The human use of human beings: Cybernetics and society*. Reissued 1989 edn. Free Association Books London.

Wiener, Norbert. 1961. *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.

Wilson, Anne, & Ross, Michael. 2003. The identity function of autobiographical memory: Time is on our side. *Memory*, **11**(2), 137–149.

Winner, Langdon. 1989. *The whale and the reactor: A search for limits in an age of high technology*. University of Chicago Press.

Wittkower, DE. 2014. Facebook and dramauthentic identity: A post-Goffmanian theory of identity performance on SNS. *First Monday*, **19**(4).

Wolf, Christopher. 2014. Impact of the CJEU's Right to Be Forgotten: Decision on Search Engines and other Service Providers in Europe: Case C-131/12 Google v. Agencia Española de Protectión de Datos (AEPO) and Mario Costeja Gonzalez, Judgment of 13 May 2014. *Maastricht Journal of European and Comparative Law*, **21**(3), 547–554.

Wolf, Maryanne, & Stoodley, Catherine J. 2008. *Proust and the squid: The story and science of the reading brain*. Icon Cambridge.

Xanthoulis, Napoleon. 2013. The right to oblivion in the information age: a human-rights based approach. *US-China L. Rev.*, **10**, 84.

Yaish, Haya. 2019. Forget Me, Forget Me Not: Elements of Erasure to Determine the Sufficiency of a GDPR Article 17 Request. *Journal of Law, Technology, & the Internet*, **10**(1), 1.

Yee, Nick, & Bailenson, Jeremy N. 2009. The difference between being and seeing: The relative contribution of self-perception and priming to behavioral changes via digital self-representation. *Media Psychology*, **12**(2), 195–209.

Yin, Robert K. 2009. Case study research design and methods fourth edition. *Applied social research methods series*, **5**.

You, Young Gweon. 1997. Shame and guilt mechanisms in East Asian culture. *Journal of pastoral Care*, **51**(1), 57–64.

Youm, Kyu Ho, & Park, Ahran. 2016. The "Right to Be Forgotten" in European Union Law: Data Protection Balanced With Free Speech? *Journalism & mass communication quarterly*, **93**(2), 273–295.

Zeller, Bruno, Trakman, Leon, Walters, Robert, & Rosadi, Sinta Dewi. 2019. The Right to be Forgotten—The EU and Asia Pacific Experience (Australia, Indonesia, Japan and Singapore). *European Human Rights Law Review*, **23**(UNSW Law Research Paper No. 19-2).

Zimmer, Michael. 2008. The gaze of the perfect search engine: Google as an infrastructure of dataveillance. *Pages 77–99 of: Web search*. Springer.

Zins, Chaim. 2007. Conceptual approaches for defining data, information, and knowledge. *Journal of the American society for information science and technology*, **58**(4), 479–493.

Zivnuska, Suzanne, Carlson, John R, Carlson, Dawn S, Harris, Ranida B, & Harris, Kenneth J. 2019. Social media addiction and social media reactions: The implications for job performance. *The Journal of social psychology*, 1–15.

Zuboff, Shoshana. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, **30**(1), 75–89.

Zuiderveen Borgesius, Frederik. 2016. Het 'Right to Be Forgotten' en bijzondere persoonsgegevens. *Computerrecht*, **126**(4), 220–225.

Zuiderveen Borgesius, Frederik J, Kruikemeier, Sanne, Boerman, Sophie C, & Helberger, Natali. 2017. Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation. *Eur. Data Prot. L. Rev.*, **3**, 353.

Zwenne, Gerrit-Jan, *et al.* 2015. Het internetvergeetrecht. *Ars*, **64**, 9–17.

Zwick, Detlev, & Dholakia, Nikhilesh. 2004. Whose identity is it anyway? Consumer representation in the age of database marketing. *Journal of Macromarketing*, **24**(1), 31–43.

Case Law

European Court of Human Rights

ECtHR, 26-04-1979, application no. 6538/74 (*Sunday Times v. UK*).
ECtHR, 05-05-1979, application no. 7805/77 (*X and Church of Scientology v. Sweden*).
ECtHR, 11-10-1979, application no. 8348/78 & 8406/78 (*J. Glimmerveen and J. Hagenbeek v. the Netherlands*).
ECtHR, 08-07-1986, application no. 9815/82 (*Lingens v. Austria*).
ECtHR, 25-06-2004, application no. 59320/00 (*Von Hannover v. Germany*).
ECtHR, 07-02-2012, application no. 3995/08 (*Axel Springer AG v. Germany*).
ECtHR, 07-02-2012, application no. 40660/08 and 60641/08 (*Von Hannover v. Germany No. 2*).
ECtHR, 13-11-2012, application no. 24029/07 (*M.M. v. the United Kingdom*).
ECtHR, 28-06-2018, application no. 60798/10 and 65599/10 (*M.L. and W.W. v. Germany*).

Court of Justice of the European Union

CJEU, 06-11-2003, C-101/01, ECLI:EU:C:2003:596 (*Criminal proceedings against Bodil Lindqvist*).
CJEU, 13-05-2014, C-131/12, ECLI:EU:C:2014:317 (*Google Spain SL, Google Inc./AEPD, G*).
CJEU, 11-12-2014, C-212/13, ECLI:EU:C:2014:2428 (*František Ryneš*).
CJEU, 24-09-2019, C-507/17, ECLI:EU:C:2019:772 (*Google v. CNIL*).

The Netherlands

Rechtbank Amsterdam, 13-02-2015, ECLI:NL:RBAMS:2015:716.
Rechtbank Amsterdam, 07-01-2016, ECLI:NL:RBAMS:2015:9515.
Rechtbank Overijssel, 25-01-2017, ECLI:NL:RBOVE:2017:278.
Rechtbank Amsterdam, 19-07-2018, ECLI:NL:RBAMS:2018:8606.
Rechtbank Midden-Nederland, 20-11-2018, ECLI:NL:RBMNE:2018:5594.
Rechtbank Limburg, 22-03-2018, ECLI:NL:RBLIM:2018:2751.

Gerechtshof Amsterdam, 31-03-2015, ECLI:NL:GHAMS:2015:1123.

Hoge Raad, 18-01-2008, ECLI:NL:HR:2008:BB3210.

Germany

Landgericht Berlin, 30-05-2013, Nr. 27 O 632/12.

Belgium

Cour de cassation de Belgique, 29-04-2016, C.15.0052.F/1.

Appendix A

BBC cases

A.1 Cases

In this appendix I provide a general overview of the BBC-articles of which the URL was delisted in Google Search in response to a name query. The cases stretch across various topics from murder to a 'war' between cooks making the best hummus. Roughly one third of the cases involve given opinions or personal experiences that were likely willingly shared by the data subject. For an overview of the covered topics, I included all article titles, divided in a list of topics that relate to illegal actions and content that relates to legal actions:

Crimes and misdemeanours - 84 cases, part 1	
Case	**Convicted, found guilty, or settled**
Man cleared of stabbing Celtic fan near Downing Street	no
Priest banned from naked calendar	yes
Three appear on explosives charges	–
Police chief 'changed sides for money'	–
Briton jailed over 'joke' e-mail	yes
PC tells corruption trial of attempted shooting	–
Charges after smuggling operation Diesel pump	–
MP's son admits theft charge	yes
Ex-TV Gladiator and detective jailed	yes
Gun-wielding 'show off' is jailed	yes
Doctor accused over Internet advice	–
Father and son jailed for 'air rage' attack	yes
Ban for drink-drive officer	yes
Pool death man took drugs cocktail	no
Hoarder' kept stolen cash in carrier bags	yes
Policeman stole 'dummy' drugs	–
Officer's 'distress' at child porn charges	no
Church settles Tolkien abuse claim	yes
From convent girl to vice queen	yes
Hotel hostage taker jailed	yes
Teacher jailed for abduction	yes
Pilot charged over crashed car	–
Briton guilty of running vice ring	yes
Madam judgment causes mayhem	yes
Jail for internet identity fraud	yes
Attack teacher keeps his job	yes
MP calls for police inquiry	yes
Officer cleared of child porn charges	no
Call-girl obsessed boss is jailed	yes
'Christian confession' over graves crime	yes
Boss's wife 'forced PA charges'	–
Former TV man guilty of assault	yes
Idiot' car thief avoids prison	no
Race case' Briton freed from jail	overturned
Jail for 'minding' dealer's drugs	yes
Two men jailed after fumes death	yes
Candidate denies illegal status	–
Binge drinker' jailed for murder	yes
Gang rape teenager was filmed'	no
Rape victim denies 'sex fantasy'	no
Footballers cleared of teen rape	no
Hacker cleared over abuse message	no
Men 'duped with date-rape drug'	yes
Rohypnol theft woman found guilty	yes

Crimes and misdemeanours - part 2	
Case	**Convicted, found guilty, or settled**
Anthrax hoax followed eBay deal	yes
Man on probation for racist jibes	yes
Plumber fined for taking a leak	yes
Force faces discrimination claims	–
Heiress killed in fit of jealousy	yes
Ex-officer loses homophobia case	no
Life term for killing ex-partner	yes
Judges decide over ball game case	no
Students jailed for train arson	yes
Casting director cleared by jury	no
Shooting linked to loyalist feud	–
Cleared of toddler's murder	no
Officer 'struck man with baton'	no
Alleged baton victim lied to jury	no
Former officer cleared of beating	no
Bomb-making kit charges dropped	no
Man refuses death crash questions	no
Men charged over 'Tigers' support	–
Policeman cleared of rape charges	no
Nanny 'caused injuries to baby'	yes
Nanny jailed for assault on baby	yes
Convicted Gladiator loses appeal	yes
Student 'crashed car into steps'	yes
Ambulance chiefs quit after probe	–
Publisher cleared of embezzlement	no
Two are guilty of insider dealing	yes
Post office embezzler avoids jail	yes
Council education chiefs probed	–
Are insider dealers afraid of the FSA?	yes
Surrey detective charged with assaulting police officer	–
Cyprus court jails drug tourists	yes
Operation Captura targets UK crime suspects in Spain	–
Women deny running illegal pyramid scheme	–
Men cleared over 'joke' Facebook looting post during riots	no
King of Marbella' John Disley jailed over bank fraud	yes
Man jailed for raping woman as she slept at house in Livingston	yes
GP injected wife with heroin at their Edinburgh home	–
Tayside Police officer 'abused ex-wife'	–
Ceredigion golf club victim locked up for days, court hears	yes
Missing girl's family jailed for not revealing her location	yes

Remaining cases - 66 cases, part I	
Case	**Likely cooperation subject**
Fears of neo-Nazi return to World Cup	–
Big Brother 2: Your views	yes
Horses die in farm blaze	–
Gran Turismo 3: Your views	yes
Sick children keep up online	yes
England's greatest ever win?	yes
Website traps speed cameras	–
Fears for missing woman	–
'Bimbo' nurse resumes army fight	–
'My student debt will top 26,000'	yes
Rape law change welcomed	–
Sue Lloyd Roberts quizzed	yes
Facial palsy left me 'isolated and bullied'	yes
Soul sold for less than 12	–
Doctor accused over Internet advice	–
Visa row for granny's new husband	–
Lost dog' dispute resolved	–
Should Vieira take a break?	yes
Out of the Bavarian backwoods	–
9.2m for road crash victim	–
Court reduces wife's pay-out	–
Why world's taps are running dry	yes
A long, hard and painful process	yes
Church settles Tolkien abuse claim	–
Schools in row over Ritalin	yes
Peak viewing for ghouls	–
Should Islamic headscarves be banned in schools?	yes
Teacher's HIV sack claim rejected	–
Seven years of struggle	yes
Is animal testing justified?	yes
In pictures: Candlelit vigil for Arafat	yes
Asian quake - Missing persons	–

Remaining cases - part 2	
Case	**Likely cooperation data subject**
How much do you value your rights?	yes
ClickBack	yes
ClickBack	yes
Former driver tells of train safety fears	yes
Views from the neighbourhood	yes
Your views: Global terror threat	yes
Asylum seeker can stay with lover	–
'Too many doctors don't know what to do'	yes
I didn't know my heart was fading'	yes
Pc expenses 'race claim' settled	–
Growing concern for missing man	–
Gay rights in the pulpit	–
Infatuated' student harassed Greer	–
Brilliant' news for lesbian couples	yes
Jerusalem Diary: Hummus wars	yes
Church to evict vicar	–
Newspaper targeted in 'evil' computer plot	–
Want help rooting our your ancestors?	yes
What it is like to live with HIV	yes
BLLCKS - (...) is a TV presenter	yes
BLLCKS - (...) is an editor	yes
Two weeks stacking shelves....	yes
Ball heading towards goal	yes
BLLCKS - (...) is a writer	yes
Merrill's mess	–
Read the latest match reports for the Mid-Wilts Youth	yes
BLLCKS - (...) is a new media salesman	yes
BLLCKS - (...) is a radiologist	yes
Ask your questions to Alan Shearer	yes
Shock tactics	yes
spooks	yes
Missing girl last seen in Glasgow	–
Woman dies in two-car crash on A482	–
Real witch Q&A	yes

A.2 Date of origin

Ninety-six of hundred-fifty the hyperlinks refer to articles older than 2007. The big number of search results pointing to relatively 'old' signifying objects is likely a reflection of Google's removal policy. However, the twelve cases referring to articles from 2014 show that people also wanted to be disconnected from relatively fresh information.

year of origin	**number of cases**
1997	1
1998	7
1999	3
2000	6
2001	10
2002	16
2003	14
2004	21
2005	10
2006	8
2007	9
2008	13
2009	4
2010	5
2011	4
2012	4
2013	3
2014	12