scripted

Volume 8, Issue 3, December 2011

FORGETTING FOOTPRINTS, SHUNNING SHADOWS. A CRITICAL ANALYSIS OF THE "RIGHT TO BE FORGOTTEN" IN BIG DATA PRACTICE

Bert-Jaap Koops[*]

Abstract

The so-called "right to be forgotten" has been put firmly on the agenda, both of academia and of policy. Although the idea is intuitive and appealing, the legal form and practical implications of a right to be forgotten have hardly been analysed so far. This contribution aims to critically assess what a right to be forgotten could or should entail in practice. It outlines the current socio-technical context as one of Big Data, in which massive data collections are created and mined for many purposes. Big Data involves not only individuals' digital footprints (data they themselves leave behind) but, perhaps more importantly, also individuals' data shadows (information about them generated by others). And contrary to physical footprints and shadows, their digital counterparts are not ephemeral but persistent. This presents particular challenges for the right to be forgotten, which are discussed in the form of three key questions. Against whom can the right be invoked? When and why can the right be invoked? And how can the right be effected? Advocates of a right to be forgotten must clarify which conceptualisation of such a right they favour – a comprehensive, user-control-based right to have data deleted in due time, or a narrower, context-specific right to a "clean slate" – and how they think the considerable obstacles presented in this paper can be overcome, if people are really to be enabled to have their digital footprints forgotten and to shun their data shadows.

DOI: 10.2966/scrip. 080311.229

[*] Professor of Regulation and Technology, TILT – Tilburg Institute of Law, Technology, and Society, Tilburg University, The Netherlands. The research leading to this article has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216483 for the project PrimeLife.

1. Introduction

Over the past years, the "right to be forgotten" has been the object of increasing attention and concern. As people realise how relentless the iron memory of the Internet can be, and how suddenly data from the past can re-emerge in unexpected contexts, many get a distinct feeling of unease. Do we want everything we say and do and look like to be recorded and stored for an indeterminate future? Should we not have a capacity, technically and legally, to have data removed if we no longer want them to roam around the Internet?

The right to be forgotten has been put firmly on the EU policy agenda and is often discussed in the literature.[1] Although reference to a right to be forgotten can already be found in the 1990s, as part of the "admirable products of European thinking and lawmaking",[2] most current literature discusses why such a right needs to be established and which possible means could be devised to effect it. However, there is no consensus what exactly a right to be forgotten means, and its status – as a right, interest, or value; in need of reinforcement or to be created from scratch – is unclear. Moreover, although the general idea is intuitive and seems widely appreciated, the legal form and practical implications of a right to be forgotten have hardly yet been analysed.

This contribution aims to assess what a right to be forgotten could or should entail in practice critically. I will focus particularly on the right to be forgotten as a legal right (rather than, for example, an abstract value), since as a lawyer I am interested in how the legal status of such a "right" can be envisioned. I will also limit myself to the European context, where the right to be forgotten has the most visible policy momentum as well as good starting points for "forgetting data" in the Data Protection Directive.[3]

This paper is structured as follows. I will first describe how the current literature conceptualises the right to be forgotten and what it would entail in practice (section 2). To appreciate more fully the challenges that the right will have to address, the current socio-technical context is sketched as one of Big Data, which consists of an accumulation of two types of data: digital footprints, i.e., data created by users themselves, and data shadows, i.e., data generated about users by others.[4] And in contrast to physical footprints and shadows, their digital counterparts are far from

---

[1] V Reding, "The Upcoming Data Protection Reform for the European Union" (2011) 1 *International Data Privacy Law* 3-5; European Commission, *A Comprehensive Approach on Personal Data Protection in the European Union* (2010); F Werro, "The Right to Inform v the Right to be Forgotten: A Transatlantic Clash" in A Colombi Ciacchi, C Godt, P Rott and LJ Smith (eds), *Haftungsbereich im dritten Millennium / Liability in the Third Millennium* (Baden-Baden: Nomos, 2009) 285-300.

[2] DH Flaherty, "Controlling Surveillance: Can Privacy Protection Be Made Effective?" in PE Agre and M Rotenberg (eds), *Technology and Privacy: The New Landscape* (Cambridge, MA /London: The MIT Press, 1998) 167-192, at 172.

[3] *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*, OJ [1995] L281/31.

[4] The distinction between digital footprints and data shadows is helpfully made by IDC, *The Digital Universe Decade* (2010) available at http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm (accessed 1 Nov 2011).

fleeting and ephemeral (section 3). I will then discuss the challenges for a right to be forgotten in the form of three key questions. Against whom can the right be invoked? When and why can the right be invoked? And how can the right be effected? These questions are discussed for three different guises of the right that feature in the literature: a right to have data deleted in due time (section 4.1), a claim on a clean slate (section 4.2), and the right to unrestrained individual expression here and now (section 4.3). The conclusion will compare the different conceptualisations and provides an outlook for further discussing and analysing the right to be forgotten (section 5).

2. The Right To Be Forgotten

2.1. Conceptual Issues

Let us briefly look at the two components of the term "right to be forgotten". First, although it is often proposed as a right,[5] some authors frame it rather as an ethical or social value,[6] or as a virtue or policy aim.[7] Rouvroy uses the interesting formulation of "*a 'right' or rather a 'legitimate interest* to forget and to be forgotten'".[8] Thus, although it may be conceived as a legal right (*de lege lata* or *de lege ferenda*), it can also be seen as a value or interest worthy of protection or a policy goal to be achieved by some means or other, whether through law or through other regulatory mechanisms.

In Rouvroy's formulation, we also see the second element expanded: she mentions not only the relevance of being forgotten but also of forgetting. Whereas the right to be forgotten uses the perspective of third parties (who should forget about your past), the right to forget uses the first-person perspective: it is also important to be able to forget your own past. This is not primarily meant psychologically (since forgetting is generally presented as a natural function of the human brain, which does not need reinforcement as such),[9] but rather has social and legal implications: the right not to be confronted with your past (which you had forgotten, or would like to forget, yourself). A convenient umbrella term for both elements – being forgotten and

---

[5] V Reding, see note 1 above, at 3-5; C Conley "The Right to Delete", *AAAI Spring Symposium Series. Intelligent Information Privacy Management* (2010); European Commission, *A Comprehensive Approach on Personal Data Protection in the European Union* (2010), at 8.

[6] J-F Blanchette and DG Johnson, "Data Retention and the Panoptic Society: The Social benefits of Forgetfulness" (2002) 18 *The Information Society* 33-45 (discussing the "social value of forgetfulness"); M Dodge and R Kitchin, "'Outlines of a World Coming into Existence': Pervasive Computing and the Ethics of Forgetting" (2007) 34 *Environment and Planning B: Planning and Design* 431-445.

[7] V Mayer-Schönberger, *Delete: The Virtue of Forgetting in the Digital Age* (Princeton: Princeton University Press, 2009).

[8] A Rouvroy, *Réinventer l'Art d'Oublier et de se Faire Oublier dans la Société de l'Information? Version augmentée* (2008) available at http://works.bepress.com/antoinette_rouvroy/5 (accessed 1 Nov 2011), at 25 (my translation; emphasis added).

[9] The psychological implications of not being able to forget are described illuminatingly in a rare case of someone suffering from hyperthymesia (superior autobiographical memory), see ES Parkeret al, "A Case of Unusual Autobiographical Remembering" (2006) 12 *Neurocase* 35-49, briefly summarised in V Mayer-Schönberger, see note 7 above, at 12-13.

forgetting – is forgetfulness[10] or the French *oubli*;[11] indeed, the concept is typically denoted in French as the *droit à l'oubli*.

Rouvroy further describes the concept as the "interest in being forgotten and in making oneself be forgotten",[12] thus making a distinction between the desired effect (being forgotten) and the possible means (ensuring that others forget). The first seems more closely associated with a negative right (a duty on others to abstain from remembering someone's past), while the latter is more closely associated with a subjective right, highlighting the liberty or power of the individual to actively control her past. The latter is also featured in the way Conley formulates his proposed "right to delete", namely that "an individual should have the right to delete information about her that is held by others".[13]

Altogether, although the details of authors' conceptualisations vary, there seems to be a considerable common denominator in the literature about a "right to be forgotten", namely that someone has a significant interest (possibly to be protected in the form of a legal right) in not being confronted by others with elements of her past, more in particular with data from the (more remote) past that are not relevant for present-day decisions or views about her.

2.2. Policy and Academic Perspectives

To flesh out in more detail what a right to be forgotten entails – and I will henceforth focus particularly on the conceptualisation as a legal right – we must look at how the current policy and literature presents this right. This, however, turns out to be surprisingly sparse. The major policy proposal is European Commissioner Viviane Reding's mention of the right as an element of the review of the Data Protection Directive (95/46/EC), which envisions

> strengthening the so-called "right to be forgotten", ie the right of individuals to have their data fully removed when they are no longer needed for the purposes for which they were collected or when he or she withdraws consent or when the storage period consented to has expired.[14]

It is telling that Reding uses the formulation "strengthening", implying that the right to be forgotten exists and is in need of reinforcement. Thus conceptualised, it seems that she understands the right to be forgotten to be nothing more or less than the

---

[10] As used by J-F Blanchette and DG Johnson, see note 6 above.

[11] As used by A Rouvroy, see note 8 above.

[12] *Ibid,* 25.

[13] C Conley, see note 5 above, at 53.

[14] V Reding, see note 1 above, at 4. This is a slight variation from the formulation in the earlier EC Communication, European Commission, *A Comprehensive Approach on Personal Data Protection in the European Union* (2010), at 8 (describing the right as "the right of individuals to have their data no longer processed and deleted when they are no longer needed for legitimate purposes").

current obligations in data-protection law to delete personal data when no longer relevant or inaccurate, or following a justified objection by the data subject.[15]

The focus on data-deletion obligations is largely in line with Victor Mayer-Schönberger's vision, who has presented the most comprehensive discussion of the right to be forgotten in academic literature to date. His main proposal to recalibrate the shifting balance between memory and forgetting, in addition to existing mechanisms that foster forgetting, is to "introduce the concept of forgetting in the digital age through expiration dates for information".[16] The same stress on the right to be forgotten consisting of deletion of old or irrelevant data can be found in much literature.[17]

However, two other approaches are visible in the literature that put forgetfulness in a slightly different perspective. The first emphasises the link with the "clean slate" or "fresh start" that has long been an element of several areas of law to foster social forgetfulness, such as bankruptcy law, juvenile criminal law, and credit reporting.[18] Similarly, Werro conceptualises the right to be forgotten as a part of personality rights that, in Swiss law, ensures that someone can preclude others from identifying her in relation to her criminal past. This focuses not so much on deletion of data, but rather on regulating the use of data:

> under Swiss law, publishing the name of someone with a criminal record may be allowed after time has elapsed since conviction only if the information remains newsworthy....This is to say that privacy concerns might preclude the press from revealing certain true and previously-publicized facts. However, the right to be forgotten will not always prevail. When information about the past is needed to protect the public today, there will be no right to be forgotten.[19]

The second alternative approach mirrors the first, in that it looks at the "clean slate" not from the perspective of society but from the perspective of the individual. Rouvroy's main concern with the right to forgetfulness is that an individual should be able to speak and write freely, without the shadow of what you express being used in the future against you. Self-development implies that you should not be fixed by what you express but that you are always free to change; the right to be forgotten thus implies the sense of liberty of expressing yourself freely in the here and now without fear that this might be used against you in the future.[20] Bannon expresses a similar concern when he states that the "collection and storing of data about peoples' activities must not be uncritically accepted. Perhaps there is a need to re-kindle

---

[15] Directive 95/46/EC, articles 6(1)(e), 12(b), and 14; see s 4.1.2 below for a discussion.

[16] V Mayer-Schönberger, see note 7 above, at 198.

[17] See e.g. C Conley, see note 5 above; J-F Blanchette and DG Johnson, see note 6 above (arguing that "control is not only a question of who has and who does not have access to personal information..., but who gets to retain or discard it"); F Werro, see note 1 above, at 285 ("privacy advocates in Europe have argued that internet users should have the right to control and possibly erase the information they leave behind themselves on the web, calling for a 'right to be forgotten'").

[18] J-F Blanchette and DG Johnson, see note 1 above.

[19] F Werro, see note 1 above, at 291.

[20] A Rouvroy, see note 8 above, at 25-26.

respect for the moment—for *being-here-now*."[21] In this approach, the focus is not necessarily only on data being deleted after their expiry date, but on a wider range of strategies that resemble the art of human forgetting. Dodge and Kitchin, writing about the rise of "life-logging" through pervasive computing, i.e. the comprehensive and continuous logging by individuals of all autobiographical events, suggest that the inexorable capacity of digital memory should be enhanced with features designed to build-in the various forms of human forgetting.[22]

Altogether, we can discern in policy and academic literature three perspectives on the right to be forgotten: a dominant perspective stressing that personal data should be deleted in due time, and two minority "clean-slate" visions: a social perspective that outdated negative information should not be used against people, and an individual self-development perspective that people should feel unrestrained in expressing themselves in the here and now, without fear of future consequences. These perspectives are not mutually exclusive, of course, but they do provide interestingly different nuances, which may become important when we want to make the right to be forgotten operational. Before I turn to that issue, let us first have a closer look at the current socio-technical context in which the right to be forgotten is supposed to take shape.

3. Socio-Technical Context

The total amount of yearly created digital information reached 800,000 petabytes in 2009 (roughly, a stack of DVDs from Earth to the moon and back) and was expected to reach 1.2 zettabytes (i.e. 10^{21} or 1,000,000,000,000,000,000,000 bytes) in 2010.[23] A lot of this is created by individuals themselves, following the rise of Web 2.0 in which every web user is also a content creator. Over 500 million people are active on Facebook, who create on average three pieces of content per day (links, stories, blog posts, notes, photos); every month, over 30 billion pieces of content are shared among users.[24] The number of people blogging online is hard to determine, but estimates vary from hundreds of thousands to several millions.[25] Blogging seems to be declining, however, now that Twitter is fast increasing, with 460,000 new accounts opened each day, with some 140 million tweets being sent each day.[26] On an average day, 43% of adult American Internet users use a social networking site, 5% upload photos, 4% create work on their own journal or blog, 4% create or work on web pages for others, and 4% share something online that they created themselves.[27] The creating and sharing of information about one-self is reinforced by trends such as

---

[21] LJ Bannon, "Forgetting as a Feature, not a Bug: The Duality of Memory and Implications for Ubiquitous Computing" (2006) 2 *CoDesign* 3-15, at 12.

[22] M Dodge and R Kitchin, see note 6 above, at 442.

[23] IDC, *The Digital Universe Decade* (2010) available at http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm (accessed 1 Nov 2011).

[24] Facebook Statistics, http://www.facebook.com/press/info.php?statistics (accessed 13 May 2011).

[25] Caslon Analytics Blogging, http://www.caslon.com.au/weblogprofile1.htm (accessed 1 Nov 2011).

[26] Twitter Blog #numbers, 14 March 2011, http://blog.twitter.com/2011/03/numbers.html (accessed 1 Nov 2011).

[27] PEW Internet Trend Data, http://www.pewinternet.org/Static-Pages/Trend-Data/Online-Activities-Daily.aspx (accessed 1 Nov 2011).

digital exhibitionism[28] and "life-logging" (wearing sensors and computers to capture everything that happens to in your life),[29] facilitated by technologies such as Microsoft's MyLifeBits.[30] Clearly, the digital footprint, i.e. the digital traces that people actively produce, is huge indeed.

Yet another big part of the zettabytes of information is not produced by web users themselves, however, at least not actively and knowingly.[31] Many data are produced about people by other parties, primarily by public and private organisations collecting and storing data about individuals in databases. A report commissioned by the Dutch Data Protection Authority estimated that the average Dutch citizen is included in 250-500 databases or in up to 1000 databases for more socially active people.[32] A large number of databases today are used for public-policy purposes. For example, in the sphere of security and justice, the EU has eighteen major initiatives and large-scale database systems involving millions of people and data-processing operations.[33] In the US alone, there are 2000 police databases.[34] Databases in the sphere of social security and social policy are similarly prevalent.

Also the private sector collects huge amounts of information. Google stores all individual search queries, not for an indeterminate period, as it did until 2007, but still for a considerable period of time, and moreover they are able to profile web users in great detail: "literally, Google knows more about us than we can remember ourselves."[35] Facebook collects huge amounts of data about people's preferences through cookies, not only of Facebook users themselves but also of non-members who simply visit a page that contains Facebook's "Like this" button, even without clicking the button.[36] Mobile phones continuously generate location data, which are stored by European telecom providers for each communication,[37] but which may also be stored on the device itself and even downloaded on users' computers without their

---

[28] N Dholakia and D Zwick, "Privacy and Consumer Agency in the Information Age: Between Prying Profilers and Preening Webcams" (2001) 1 *Journal of Research for Consumers*, at 13.

[29] S Mannet al, "Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments" (2005) 1 *Surveillance & Society* 331-355.

[30] Microsoft Research, *MyLifeBits: A Personal Database for Everything* (2006).

[31] V Mayer-Schönberger, see note 7 above, at 88-90.

[32] Considerati, *Onze digitale schaduw. Een verkennend onderzoek naar het aantal databases waarin de gemiddelde Nederlander geregistreerd staat* (2009).

[33] EC, *EU Information Management Instruments* (2010).

[34] E Murphy, "Databases, Doctrine & Constitutional Criminal Procedure" (2010) 37 *Fordham Urban Law Journal* 803.

[35] V Mayer-Schönberger, see note 7 above, at 7.

[36] A Roosendaal, "Facebook Tracks and Traces Everyone: Like This!" (2010) *Tilburg Law School Research Paper No. 03/2011*, SSRN available at http://ssrn.com/abstract=1717563 (accessed 1 Nov 2011).

[37] *Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC (Data Retention Directive),* OJ [2006] L105/54.

knowledge.[38] To be sure, not all of these data are personal data traceable to identifiable persons, but through the vastness of the data "out there" there is always a possibility that data are combined with other data and then suddenly become personal data. Moreover, the data are also used for profiling, and as such can become relevant at any time when a profile based on anonymous data are applied to an individual in a decision or service offer.[39]

Finally, not only organisations produce data about individuals, also individuals themselves are increasingly active in generating data about other people, through for example blog posts and tweets. An illustrative activity is photo tagging: the adding of someone's name to a photo on a social networking site, so that the photo can be automatically linked to the person's page.[40] With millions of photos uploaded daily[41] and many millions of Facebook users, there is a significant likelihood that someone can be recognised and tagged.

In short, it is clear that we live in a world of Big Data. Particularly relevant for a right to be forgotten is the accumulation of data from different times and places: "What makes most big data *big* is repeated observations over time and/or space."[42] Significantly, the zettabytes of information generated over the years are not only created actively by people themselves, but also created by leaving traces that others collect or generate. According to one source, the second development has outgrown the first: more data are nowadays created *about* individuals than *by* individuals. In other words, our "digital shadow" has outgrown our "digital footprint".[43] If this is true, it has considerable implications for how we can shape a right to be forgotten.

4. Implications for a Right To Be Forgotten

Given a world of Big Data, of which digital shadows form at least as important a part as digital footprints, how could or should we conceive of a right to be forgotten? In this section, I will try and make such a right more concrete, in its three (interrelated but distinguishable) possible guises[44] as a right to have data deleted in due time (s 4.1), a claim on society to have a "clean slate" (s 4.2), and an individual interest in unrestrained expression in the here and now (s 4.3). For each incarnation, I will try and clarify three dimensions of the right (against whom, when and why, and how), illustrated by some examples.

---

[38] BBC News, "Apple 'Not Tracking' iPhone Users", 27 April 2011, http://www.bbc.co.uk/news/technology-13208867 (accessed 1 Nov 2011).

[39] Cf. M Hildebrandt, "Profiling and the Identity of the European citizen" in M Hildebrandt and S Gutwirth (eds), *Profiling the European Citizen* (s l.: Springer, 2008) 303-326.

[40] A Besmer and H Richter Lipford "Moving Beyond Untagging: Photo Privacy in a Tagged World", *CHI 2010* (2010).

[41] The record being 750 million Facebook photos uploaded in a single weekend, see Jason Kincaid, "Facebook Users Uploaded A Record 750 Million Photos Over New Year's", *TechCrunch* 3 Jan 2011, http://techcrunch.com/2011/01/03/facebook-users-uploaded-a-record-750-million-photos-over-new-years/ (accessed 1 Nov 2011).

[42] A Jacobs, "The Pathologies of Big Data" (2009) 52 *Communications of the ACM* 36-44.

[43] IDC, see note 23 above.

[44] See s 2.2 above.

4.1. Delete Data in Due Time

4.1.1. A Right Against Whom?

As a right to have data deleted in due time – which might mean: after use, when no longer relevant, when an expiry date elapses, or when the drawbacks of retention start outweighing the advantages – the right to be forgotten already seems part of the current data-protection framework. After all, data subjects have the right to see personal data deleted when they are no longer relevant or inaccurate, or following a justified objection.[45] This right can be invoked against the data controller: the one who, alone or jointly with others, determines the purposes and means of the processing of personal data.[46] This will typically be a public or private organisation, but it can also be an individual, for example, in Web 2.0 applications.

Although the concept of controller was relatively easy to apply in the 1990s when the Directive was adopted, it is becoming more difficult to determine who exactly determines the purposes and means of data processing in today's hybrid information processes. For example for mobile-phone location data, it is clear that the telecom provider is the data controller who has the obligation to delete the data after the retention period determined by law.[47] For other examples of data, however, such as blog postings and uploaded videos and photos, the situation is less clear since we are dealing here with Web 2.0, in particular social-networking services (SNS). The Article 29 Working Party, in their opinion on the concept of controller, identifies the SNS service providers as "data controllers, since they determine both the purposes and the means of the processing of such information." At the same time, the "users of such networks, uploading personal data also of third parties, would [also] qualify as controllers provided that their activities are not subject to the so-called 'household exception'".[48] Here, users are mentioned as possible data controllers if they upload information about others ("also of third parties"). This applies for example to blog posts and YouTube videos if these contain data about other identifiable individuals, and to users tagging photos of others. The "household exception" does not apply if the user "acts on behalf of a company or association, or uses the SNS mainly as a platform to advance commercial, political or charitable goals". The exception might also not apply if a user has a "high number of third party contacts, some of whom he may not actually know".[49]

For the right to be forgotten, two situations are relevant: one in which the user wants to remove data she has uploaded herself, such as the blog post or video, and one in which the user wants to have data removed uploaded by other users. In the first case, although the user might be data controller at the same time (if data about others are involved and no household exception applies), the primary data controller will be the platform provider, i.e. the blog host and YouTube. The user should be able to remove

---

[45] Directive 95/46/EC, articles 6(1)(e), 12(b), and 14; see s 4.1.2 below for a discussion.

[46] Art. 2(d) Directive 95/46/EC.

[47] *Directive 2006/24/EC*, OJ [2006] L105/54.

[48] Article 29 Data Protection Working Party, *Opinion 1/2010 on the Concepts of "Controller" and "Processor"* (2010), at 21.

[49] Article 29 Data Protection Working Party, *Opinion 5/2009 on Online Social Networking* (2009), at 6.

the content herself or request the SNS provider to remove the data. In the second case, the situation is more complex, as both the SNS provider and the user/uploader can be data controller at the same time. Suppose that Alice wants to have removed a blog post by Bob mentioning what she told him during dinner a while ago, or a photo uploaded by Bob featuring Alice and Bob embraced in a rather intoxicated pose, and tagged by Carol. Should she approach Bob with the request to remove the post or photo, Carol to remove the tag, or the SNS provider to remove either of these? Much will depend on the concrete distribution of responsibilities between provider and users defined in the terms and conditions, or developed in practice, of the concrete SNS application.

There are at least two complicating factors here. First, if Bob's blog post and Carol's tagging are subject to the household exception,[50] which can be the case with many individuals' social-networking contributions,[51] Alice has no claim on Bob or Carol to have the information removed. In those cases, it is also doubtful whether she still has a claim on the SNS provider, who despite being a data controller in its own right, might say with some justification that Bob and Carol are the primary data controllers for these particular data, and hence lay aside a request to remove the information. Alice still has the right vis-à-vis the SNS provider "to object at any time on compelling legitimate grounds relating to [her] particular situation to the processing of data relating to [her]",[52] but that requires her to explain why it is compelling for her to have the blog post or photo (or tag) removed (likely involving further personal data provision to the SNS provider), while the provider has a discretionary power to determine whether the grounds are sufficiently compelling – also taking into account Bob's and Carol's right to freedom of expression.

Second, if the household exception does not apply, the distribution of responsibilities is not particularly clear, since both the SNS provider and the user/uploaders are being designated as data controllers in the standard interpretation of the Directive.[53] Not only can this confuse Alice, but it also creates a considerable risk that the SNS provider and Bob & Carol can refer Alice to the other party, claiming that removal is not their responsibility but rather the other's.

More complications arise due to the multiplying character of Internet data. Even if data are removed at the source, copies can still be retained in caches (technical measures to foster efficiency on the net) or on mirror websites. Although users will have the same right to request removal from such secondary sites as from the primary source, and the caching or content providers have similar responsibilities as data controllers to remove the data, it will be a greater challenge for users to identify the particular providers that host copies. Caching providers can perhaps be found sufficiently through search engines and will generally have an interest in complying

---

[50] Art. 3(2) Directive 1995/46/EC ("This Directive shall not apply to the processing of personal data…by a natural person in the course of a purely personal or household activity").

[51] Contributions to SNS sites by individual users are usually not made on behalf of a company or for substantially commercial, political, or charitable purposes; only individual users with (suspiciously?) many contacts ("some of whom he may not actually know") would then have an obligation to remove data.

[52] Art. 14 Directive 1995/46/EC.

[53] See note 48 above.

with a request for removal,[54] but mirror sites may be less traceable or co-operative. There is a significant tendency on the Internet to copy material that is considered funny or embarrassing, and particularly when official requests are made to remove material at the source there are actors who, acting upon a sweeping vision of the freedom of expression, copy material to preserve it for the online community.

Besides the data controllers responsible for concrete pieces of data, there may be yet another party that data subjects could desire to approach for enacting their right to be forgotten. This is the state, more in particular the legislature, who is increasingly passing legislation obliging parties to retain (rather than delete) data. The most visible example is the European Data Retention Directive, but requirements to store and retain data are also part and parcel of much sectoral regulation.[55] There is no intrinsic reason to limit a right to be forgotten to data that are no longer relevant to keep except if a legal obligation determines they have to be stored. Alice may have equally good reasons to object to her telecom provider storing where she was on 15 July, 2011, as she may have to Flickr identifying her as the tipsy girl in Bob's arms. But if the telecom provider is legally obliged to store her location data, Alice has no claim on the provider. All she can do is go to court and claim that the legal data retention obligation violates her fundamental rights, in particular her right to privacy. Although in some cases of extreme retention periods, such a claim can be successful – as S. and Marper managed to have their DNA profiles removed from the English DNA database –,[56] it is doubtful whether individuals can in general lay a successful claim on governments to abolish or shorten legislative retention periods.[57]

In conclusion, it will often be clear against whom a right to be forgotten can be exercised, namely the data controller who has the primary responsibility for the data at issue. Nevertheless, difficulties arise in Web 2.0 situations where data you want to be deleted are uploaded by others who can often fall under the household exception, or where the fuzzy responsibility allocation between SNS provider and users makes it difficult, in theory if not in practice, to pinpoint the right target for having data deleted. Also parties responsible for copies in caches or mirror sites can be difficult to address. Moreover, for several types of data users have no claim on data controllers as these are subject to legislative obligations to retain data, and it is doubtful whether users can successfully target the state to challenge the retention of data they have an interest in seeing removed.

---

[54] Cf. art. 13(1)(e) *Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market ("Directive on electronic commerce"),* OJ [2000] L1781 (stipulating that caching providers are not liable if "the provider acts expeditiously to remove or to disable access to the information it has stored upon obtaining actual knowledge of the fact that the information at the initial source of the transmission has been removed").

[55] V Mayer-Schönberger, see note 7 above, at 160-162.

[56] ECtHR 4 December 2008, app. 30562/04 and 30566/04 (*S and Marper v The United Kingdom*).

[57] The national implementations of the Data Retention Directive have been successfully challenged on constitutional grounds in Germany and Romania, but this does not apply – so far – to the Directive itself. Also note that the German decision did not object to retention as such (considering the retention period in the German law of six months acceptable), but rather the (too wide) scope of use of the data for law-enforcement purposes. See BVerfG 11 Mar 2008, 1 BvR 256/08, http://www.bverfg.de/entscheidungen/rs20080311_1bvr025608.html (accessed 1 Nov 2011).

4.1.2. When and Why Can the Right Be Invoked?

I see two main reasons why a right to be forgotten can be invoked: because the retention of data is in any way (potentially) harmful to the individual (a concrete risk in the here and now) or because the data are no longer necessary while their retention might become harmful to the individual at some point (an abstract risk in the future).

The first case seems relatively straightforward. It is partly regulated in the current Data Protection Directive in art. 14:

> Member States shall grant the data subject the right...to object at any time on compelling legitimate grounds relating to his particular situation to the processing of data relating to him, save where otherwise provided by national legislation. Where there is a justified objection, the processing instigated by the controller may no longer involve those data....

Although this right can be invoked at all times, it suffers from some limitations. As already noted, it requires data subjects to substantiate that there are compelling legitimate grounds to stop data processing, which puts a significant burden of proof on users and leaves large discretionary power with the data controller.[58] Moreover, the objection need not necessarily be considered justified for all *forms* of processing (which includes, *inter alia*, collection, storage, adaption, use, and disclosure).[59] The controller could argue, for example, that a data subject has a justified objection to the data being used or disclosed, but not to storage of the data, as this is less privacy-infringing. There could remain then a residual risk that the data are used in the future when the data controller thinks this is warranted, and then the data subject would have to object to the data processing again and again. Finally, the provision provides an exception for legal obligations to process the data. Although some such exception is needed (as governments and business must have some power to process data against the interest of data subjects themselves, e.g., to combat crime or fraud), it leaves a broad discretion with national legislatures to make exceptions to individuals' interest in objecting to data processing.

Another instantiation of the right to be forgotten protecting against a concrete risk of harm seems to be art. 12(b) of the Data Protection Directive:

> Member States shall guarantee every data subject the right to obtain from the controller (...) as appropriate the rectification, erasure or blocking of data the processing of which does not comply with the provisions of this Directive, in particular because of the incomplete or inaccurate nature of the data (…).

This provides users with a stronger right, in that it enables them to request erasure, without a generic exemption for legal obligations providing otherwise; there is also no need to extensively substantiate the reason for desiring data to be deleted, and the data controller has less room for interpreting the justification of the request. However, this

---

[58] See note 52 above and surrounding text.

[59] See art. 2(b) *Directive 95/46/EC*, OJ [1995] L281/31.

situation is restricted to data processing that violates the terms of the Directive, in particular when incomplete or inaccurate data are processed. The scope can be broad if "in particular" is interpreted as non-limitative, which may depend on the exact implementation of this provision in national legislation. The provision, in a generous, data-subject-friendly reading, then contains a general right to request erasure of data that are processed unlawfully. A less generous reading is also possible, if the right to request erasure is interpreted "particularly" by concentrating on incorrect data. In any case, whether the scope of the provision is interpreted more or less broadly, art. 12(b) does not boil down to a general right to be forgotten, which also needs to address correct and lawfully processed data. Art. 12(b) gives Alice the right to address Bob, Carol or the SNS provider when Carol misidentified the tipsy girl in the picture as Alice (while in fact it was Annie), but does it give a title to have the photo or tag removed if indeed it was Alice? Perhaps it does, but this puts a burden of proof on Alice to show that the uploading or photo tagging was unlawful, instead of Alice having a right to request deletion simply because she has a prima facie interest to see potentially harmful data forgotten. There are also other legal rights that can be invoked in cases of correct but harmful data, in particular personality rights (such as portrait rights) or defamation, but these also require considerable effort to invoke in terms of substantiating the claim and convincing the data controller, or the court, that the harm outweighs other interests such as freedom of expression.

Although there are altogether several legal grounds to request removal or at least stopping the processing of data that are harmful to an individual, there is a major drawback in invoking the right to be forgotten in these cases. Typically, the harm will already have materialised: because the rights work *ex post*, the data subject will usually only become active after she has noticed that detrimental data are being processed. Although Alice may of course be reading Bob's blog posts daily and do vanity searches on the Internet to see whether new information has been posted about her, often she may only notice the harmful blog or picture when it is too late: when she is denied the promotion she hoped for, when customs do not let her enter the United States because of that confession of teenage drug abuse she blurted out over dinner with blogger Bob, or now that a surprising number of ads about alcohol-addiction treatments show up on her screen. She might be able to invoke the right to be forgotten to have the notorious data removed, thus preventing possible further harm in the future, but that does little to redress harm already done.

That is why the second reason is particularly important: users have an interest in seeing data deleted when they are no longer relevant, to prevent the data from lingering around and coming back somewhere in the future with a vengeance. This is obliquely addressed by art. 6(1)(e) of the Data Protection Directive:

> Member States shall provide that personal data must be (...) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use.

This contains a general obligation on data controllers to delete data after use (or to anonymise data, which amounts, in principle,[60] to the same thing for the purposes of individuals' right to be forgotten). Although this is an obligation on controllers rather than a right for data subjects, we could perhaps read a right to be forgotten into this in connection with art. 12(b) discussed above, since data subjects can request erasure if controllers violate art. 6(1)(e) by storing data longer than necessary (although only if we read the subordinate clause "in particular because of the incomplete or inaccurate nature of the data" as non-limitative). But if we are concerned with preventing harm, the same objection applies here as with art. 6(1) if the deletion obligation is only enforced when data subjects complain of violations: the harm will often already be done.

Therefore, the deletion of irrelevant data has a crucial prospective aspect: a good governance system must be in place that ensures data controllers delete data immediately after they have served their purpose. To guarantee this, Mayer-Schönberger proposes that users should systematically set expiry dates for data, on top of existing data-protection rights and duties.[61] Through the automatic deletion of data after their expiry date, this would ensure that users can exercise their right to be forgotten ex ante: before generating, storing, or sharing data, without having to worry about how long the data will linger around without their knowledge or consent.

Attractive though it might sound to thus shape a right to be forgotten, some fundamental problems are associated with systematically setting expiry dates. First, users will have to make decisions on how long they think they will need the data to be processed. For some types of data, for example data for a concrete, contemporaneous context such as tweets or subsidy requests, this will be relatively easy, but for many types of data, users will have difficulty in foreseeing future use(s) or usefulness of the data. How long do you want a photograph to be on your Facebook page? How long do you think your blog will be interesting to read? Mayer-Schönberger argues that the point of having to set expiry dates is precisely forcing users to think about these questions,[62] but I am sceptical that users would really reflect and make informed decisions about expiry dates. Many could simply set "infinity" as a default for most social-networking types of information such as blogs, comments, and photos – you never know when you might want to see it back in the future. Setting expiry dates also clashes with those intent on life-logging: preserving as much of their life experiences as they can. Also convenience provides an argument for setting longer rather than shorter retention periods: if you provide data to others, it is handy if they keep a copy so that next time you don't have to fill out the entire form again.

More fundamentally problematic is that even if users were able to make informed decision about expiry dates for data concerning them, often they have no such decisional power. After all, it is the data controller, not the data subject, who defines the purposes of data processing, and it is therefore logical that the controller will determine when exactly the purpose has been fulfilled. In other words, the data

---

[60] Note, however, that through merging large databases, anonymous data could become personal data again, and through profiling, non-personal data can also reveal information about individuals; see note 39 above and surrounding text. Data subjects therefore may have an interest in having data deleted rather than anonymised.

[61] V Mayer-Schönberger, see note 7 above, at 171-195.

[62] *Ibid,* 171.

subject has a say in the purposes for which data are processed, and she can give a generic consent for processing for the duration that these purposes remain relevant, but she can hardly specify a concrete period for processing. If users should be able to do so, in line with Mayer-Schönberger's proposal, the Data Protection Directive must be adapted to enable users to set expiry dates.

Even if the Directive were changed accordingly (which I think is difficult in light of organisational interests, not only in being able to decide about retention periods themselves but also to avoid additional administrative burdens in adapting IT systems and database structures), there will still be limits to what users can do in terms of defining retention periods ex ante. As noted, statutory retention periods exist for many types of data,[63] which will override individual user preferences for expiry dates. At least equally significantly, setting expiry dates may work for digital footprints, but not for data shadows. How could a user define a retention period for data that are generated about them by others without their direct involvement or knowledge? Bob and Carol might be forced by systems to set expiry dates for the blog post, photo, and tag they put online, but they will do this based on their own interests and preferences, not on Alice's.[64]

To address the problem of automatic expiry of data shadows, we may therefore have to resort to legal obligations on third parties to set expiry dates that take into account the interest of data subjects. But one can doubt whether that would make a material difference. Such legal obligations would either have to be very detailed and sector-specific, for those contexts in which it is sufficiently clear for the legislature to determine specifically how long (or rather, how brief) retention periods would have to be, or they would be generic and take the form of procedural requirements, which would leave large discretionary powers for data controllers to make trade-offs in setting the expiry periods. Another issue is that having legally required user-friendly retention periods on paper does not necessarily mean data deletion in practice; the effectiveness hinges on the extent to which such legal obligations would be enforced.

Apart from the difficulties associated with expiry dates in dealing with both footprints and shadows, there are two other fundamental complications with exercising a right to be forgotten through having data deleted "after use", be it in the form of ex ante expiry dates or otherwise. The first is that the right in this guise can be invoked only *after* use, i.e. when the data no longer serve the purpose for which the data were collected. But harm can also occur during the period of legitimate processing, or before the expiry date set by the user has elapsed. The risk of harm may be less during legitimate and timely processing; after all, the debate about the right to be forgotten is triggered by increasingly long retention of outdated data where there is a higher risk of unfair judgement. Nevertheless, if users were to set, as I expect them to, rather longish expiry dates for their footprints, and if third parties set longer expiry dates for shadows, the risk that data cause harm to data subjects during legitimate processing may be non-negligible. A right to be forgotten focusing only on outdated data does not address this risk.

The second fundamental complication is that determining when data are no longer relevant is easier said than done. The purpose-limitation principle indicates that data

---

[63] V Mayer-Schönberger, see note 7 above, and surrounding text.

[64] C Conley, see note 5 above, at 57.

should be deleted when no longer necessary for the purposes for which the data were collected *or for which they are further processed*. Data can thus be retained if this is needed for so-called secondary use, for a purpose that is not incompatible with the initial purpose. In the socio-technical context of Big Data, secondary uses of data have become much more important than they were in the 1970s and 1980s. For businesses, direct marketing based on increasingly refined user profiles has become a primary business operation, while for government, data analysis and sharing for fraud detection and other forms of risk governance occurs on a much wider scale than one or two decades ago. Moreover, data are also increasingly being collected for yet unknown or rather vague initial purposes, following the logic of data mining that huge data sets can reveal new and unexpected knowledge. The challenge to purpose-limitation is well captured by the notion of "function creep", which indicates the situation "when a system developed for a particular purpose comes to be used for, or to provide the underpinnings for other systems that are used for, different purposes".[65] The attention in the literature to function creep since the early 1990s[66] is indicative of a growing concern in recent years over the use of data for other purposes than those for which they were originally collected.

Now that the socio-technical context of Big Data implies that data processing is being based on vague purpose definitions to allow unforeseen future uses, and that data are increasingly used for secondary purposes, this fundamentally challenges not only the purpose-limitation principle itself[67] but also the effectiveness of a right to be forgotten if that is tied to retention periods based on legitimate processing for purposes that are vague or derivative.

In conclusion, if the right to be forgotten focuses on deletion of data that are no longer relevant, this raises significant issues for when and why the right can be invoked. One problem is the trigger event: a data subject may be alerted to the existence and use of outdated data when it is too late, i.e., when harm is already done through a decision being made on the basis of the outdated data. If this is to be prevented, the right should have an ex ante operationalisation, through a governance system that ensures that data processors delete data that are no longer necessary to keep, possibly reinforced by automatic deletion of data when an expiry date set by users elapses. But there are significant obstacles to this approach: for users, it is difficult to make considered choices on expiry dates, which do not work well in any way for data shadows or for data with statutory retention periods. Moreover, determining when data are outdated in the era of Big Data is easier said than done, with function-creeping information systems erring on the side of longer retention for secondary or newly emerging purposes. Finally, it is not only outdated data that can be detrimental to data subjects: harm can also occur during legitimate processing periods, and a right to be forgotten focusing only on outdated data does not address that concern. This is alleviated somewhat by current rights to object to processing because of a compelling interest, such as art. 12 of the Directive or portrait rights, but these require substantial

---

[65] MR Curry et al, "Emergency Response Systems and the Creeping Legibility of People and Places" (2004) 20 *The Information Society* 357-369, at 362.

[66] It first occurred in R Clarke, "Tax File Number Scheme: A Case Study of Political Assurances and Function Creep" (1991) 7 *Policy* (4).

[67] BJ Koops, "The (In)flexibility of Techno-Regulation and the Case of Purpose-Binding" (2011) 5 *Legisprudence* 171-194.

effort by data subjects and moreover involve only stopping the processing, not necessarily also the deletion of data.

4.1.3. How Can the Right Be Effected?

In order to ensure that data are deleted in due time, considerable regulatory effort is required, particularly in light of the complications to the right to be forgotten that we already encountered in terms of who, when, and why data subjects can address with a request to have data deleted. Most authors focus on a combination of legal and technical regulatory measures. I will leave aside the occasional reference to social or economic measures, which are little elaborated and unrealistic.[68] To give one example:

> Societal buy-in may be both necessary and sufficient to establish a right to delete. If members of society can agree that individuals deserve the right to own their own digital persona, including records that are held by third parties, then a right to delete can be established absent any legal change…and market actors will adapt to changing consumer expectations. Of course, actors that rely on monetizing records about individuals may resist this trend, but these same actors are often susceptible to collective action and public pressure.[69]

This hinges on a very big "If" in the beginning of the quote. Absent any suggestion of how such social changes and public pressure could realistically come about in the world of Web 2.0 and Big Data,[70] relying on social or economic regulatory measures is not a feasible proposition. I will therefore restrict myself here to legal and technical measures, first discussing law and then design.

As to legal measures, the Data Protection Directive (DPD) already provides a good starting point, but it has several weaknesses in current provisions to effect a right to be forgotten. If this right is to be seriously shaped in the review of the Directive, the problems identified in the previous sections will have to be addressed, including:

- the household exception for users uploading content about others in social networking contexts;
- the fact that data subjects, during the legitimate processing period, can request stopping processing but not necessarily also erasure when they have a compelling interest, and the significant burden of showing such compelling grounds (art. 14);
- the fact that the obligation to delete data after use (art. 6(1)(e)) leaves data controllers with considerable discretionary power to define vague purposes or secondary purposes that enable them to retain data much longer than for primary short-term purposes;

---

[68] Cf Mayer-Schönberger's discussion of digital abstinence, V Mayer-Schönberger, see note 7 above, at 128-134, which, for example, assumes without substantiation that users could be educated "to choose carefully before sharing their personal information with others". *Ibid,* 132.

[69] C Conley, see note 5 above, at 58.

[70] See s 3 above.

- the limitation of requesting erasure in case of DPD violations only, or "in particular", when data are inaccurate or incomplete (art. 12(b)); this might have to be extended with a ground for objection when data are processed for vague or derivative purposes in "function-creeping" systems;
- the fact that the Directive contains no right for data subjects to set concrete expiry dates but allows determination of processing periods only indirectly through users consenting with the purpose(s) of processing, which leaves the data controller with a wide scope for interpreting how long the data are needed to fulfil the (secondary) purpose(s);
- the fact that, if data are processed on other legitimate grounds than consent – which will usually be the case with data shadows –, the interest of data subjects in setting data expiry periods is not even indirectly taken into account.

Altogether, these are huge challenges for the Data Protection Directive review to take up when shaping a right to be forgotten. Moreover, even if these points were addressed through adapted provisions in the revised Directive, it would still come down to enforcement to effect the right to be forgotten in practice. Enforcement will likely need to be stepped up, not only in terms of improved capacity, powers, and policies of Data Protection Authorities,[71] but also by more attention for privacy by design (i.e. technical reinforcement)[72] which I will discuss below.

Perhaps more radical legal measures are called for, however, since revising the Directive to embed a right to be forgotten, even if it were feasible, may not be enough. Two fundamental problems are difficult to address through data-protection legislation in its current form. The first is the issue of legitimate processing periods, during which data can also be detrimental to data subjects. Even if articles 12(b) and 14 of the DPD were revised, they would still need to contain some sort of balancing of interest between the data subject's legitimate interest in having data deleted and the data processor's legitimate interest in processing the data. Given the power imbalance that frequently exists between data subjects and data controllers,[73] this leaves data subjects in an awkward position to substantiate their compelling grounds to desire data to be deleted. This could be addressed by framing data subjects' control over their data not as (sui generis or tort-based) data-protection rights but as property rights. Although propertisation of personal data does not at first sight seem compatible with the European approach to data protection,[74] Purtova has convincingly argued that a meaningful legal debate can be held in Europe about shaping data-protection interests as property rights, through carefully distinguishing between the various elements and purposes of property.[75] The argument that Conley applies to the

---

[71] RAND, *Review of the European Data Protection Directive,* TR-710-ICO (2009), at 35-36.

[72] A Cavoukian, "Privacy by Design: The Definitive Workshop. A Foreword" (2010) 3 *Identity in the Information Society* 247-251; V Reding, see note 1 above, at 4.

[73] Cf V Mayer-Schönberger, see note 7 above, at 100-112; BJ Koops, "Law, Technology, and Shifting Power Relations" (2010) 25 *Berkeley Technology Law Journal* 973-1035.

[74] JEJ Prins, "Privacy and Property: European Perspectives and the Commodification of our Identity", in L Guibault and PB Hugenholtz (eds), *The Future of the Public Domain* (The Hague: Kluwer Law International, 2006) 223-257.

[75] N Purtova, *Property Rights in Personal Data: a European Perspective,* PhD Thesis, Tilburg University (2011).

American context, namely that "[p]roperty law provides a better model [than privacy torts], giving individuals affirmative rights without any need to demonstrate harm",[76] also holds a valid point for Europe. An additional advantage of a property approach to the right to be forgotten is that because of the "erga omnes" effect (i.e. a property right can be invoked against anyone, instead of only in bilateral legal relationships),[77] it could better address the problem of copies of data lingering around the Internet, for example on mirror sites.[78] At the same time, we must be cautious in putting our cards on propertisation for shaping a right to be forgotten, since the right to delete is the most far-reaching element of the "bundle of sticks" of property rights, which contains *usus* (use), *fructus* (reap the fruits), and *abusus* (modify or destroy).[79] Part of the attractiveness of a property solution to data-protection problems is that the legislature can treat the various sticks from the bundle differently, and not necessarily introduce all property rights for personal data.[80] However, if a right to be forgotten is cast in the form of a property right for data subjects to delete (*abusus*), this being the strongest property stick, then likely the entire bundle of sticks would have to be allocated to data subjects, which seems a bridge or two too far.

The second fundamental issue that is hard to tackle through current data-protection law is the fact that data deletion cannot be requested in case of a legal obligation to retain data. If we observe the keenness with which legislatures are passing retention obligations, not only for crime-fighting and anti-terrorism but also more generally for accountability and transparency,[81] a right for data subjects to request deletion of data may be of limited value if they cannot challenge data-retention laws in their own right. While article 8 of the European Convention on Human Rights (ECHR) gives some basis for citizens to do so, this is a cumbersome and very lengthy path for individuals to take. Moreover, it is an uncertain path as well, since legislatures have a considerable margin of appreciation to determine what they consider "necessary in a democratic society" (art. 8(2) ECHR), and since a right to be forgotten can only be read very indirectly into the right to privacy as formulated in art. 8(1) ECHR. Perhaps, then, a clearer constitutional basis is needed for data subjects' right to request data deletion. Significantly, such a right is not part of the data-protection rights as constitutionalised in the Charter of Fundamental Rights of the European Union, in which art. 8(2) stipulates that everyone "has the right of access to data which has been collected concerning him or her, and the right to have it rectified", but not to have it deleted. For a really effective right to be forgotten, art. 8 of the Charter may therefore have to be extended with a right to have data erased, so that citizens could challenge comprehensive European data-retention legislation. Amending the Charter, however, is not a realistic option, seeing the enormous political and legal difficulties the EU experienced before it came into force.

As already noted, technological measures will need to reinforce or supplement legal measures. Most proposals concern systems to ensure the automated deletion of data,

---

[76] C Conley, see note 5 above, at 55

[77] N Purtova, see note 75 above, at 239-241.

[78] See s 4.1.1 above.

[79] N Purtova, see note 75 above, at 65.

[80] *Ibid*, 256-258.

[81] See note 55 above and surrounding text.

in line with Mayer-Schönberger's proposition of expiry dates that I have already discussed in the previous section.[82] The techno-regulation (i.e. regulation by technical design) could be more refined than black-or-white retention versus destruction; Dodge and Kitchin provide interesting suggestions for measures "that should be built into the system" to resemble human forgetting in all its forms.[83]

> Just as a person would simply start to forget parts of the journey, so the life-log would gradually degrade the precision of the record with time. Absentmindedness could be ensured through distractedness being built into the sensing technologies (...). Misattribution could be achieved by the specific misrecording of part of an event, but not the whole event. For example, part of a journey would be randomly misattributed (eg, having a coffee in Starbucks rather than Caffè Nero) (...). In other words, misattribution is meaningful in the relations of time, space, and context. It is not the adding of false memories, but rather the "tweaking" of a past event. (...) Overall, then, a range of algorithmic strategies could be envisioned, such as erasing, blurring, aggregating, injecting noise, data perturbing, masking, and so on, that would be used to "upset" the life-log records.[84]

A broader perspective is also suggested by Conley, who rejects the focus on expiry dates as these work only with digital footprints and not with data shadows. Instead, he advocates "a manual ability to delete records. From the user side, the key technology needed to enable the framework we envision is a 'deletion manager'".[85] Some such tools are already being developed, such as the Web 2.0 Suicide Machine that allows users to erase records and profiles from multiple social networking sites simultaneously.[86] This approach requires not only tools enabling people to erase their footprints or data shadows, but crucially also discovery mechanisms so that data subjects will know what information about them is being processed "out there".[87] Unfortunately, Conley does not explain which tools are able to give "an individual the information she needs to decide when or whether to delete records without being overwhelmed by the volume of records that capture her life in otherwise permanent storage".[88] Indeed, "deletion managers" need to be very smart in learning which data shadows to inform a user about without causing instant shadow-exhaustion syndrome,

---

[82] V Mayer-Schönberger, see note 7 above; LJ Bannon, see note 21 above, at 12 (suggesting "various kinds of electronic tagging systems for messages or material that could time-stamp material and contain something like a 'sell-by' date, where the information would self-destruct after the elapsed time").

[83] M Dodge and R Kitchin, see note 6 above, at 442.

[84] *Ibid*.

[85] C Conley, see note 5 above, at 57.

[86] Web 2.0 Suicide Machine, http://suicidemachine.org/ (accessed 1 Nov 2011). The website advises a walk in the park or buying a bottle of wine to enjoy real life after digital suicide, http://suicidemachine.org/#faq (accessed 1 Nov 2011).

[87] C Conley, see note 5 above, at 57.

[88] *Ibid*.

and in facilitating automated deletion as much as possible without risking undesired deletion.

Although the proposals for designing in forgetting in the Internet architecture are useful, they are still largely embryonic when it comes to implementation and large-scale adoption. Much work will need to be done before comprehensive techno-regulation will facilitate forgetfulness. The notion of privacy by design currently benefits from considerable policy attention, but we are still in the stage of inventorying ingredients and recipes for forgetfulness technologies before we can even think of putting this pudding to the proof through eating. It bears remembering that the precursor of privacy by design, Privacy Enhancing Technologies (PETs), has been a similarly promising concept ever since the idea was launched in 1995,[89] but that during the past decade and a half PETs have still not been widely adopted. Considerable hurdles need to be overcome before data controllers (and data subjects) would start deploying PETs on a larger scale.[90] Particularly significant is the finding that "data controllers often favour mere data protection to protect themselves against the adverse consequences of data loss over data minimisation or consent mechanisms which can impede the use of personal data."[91] Technical solutions focusing on empowering data subjects to have their data deleted are therefore unlikely to find a warm welcome with data controllers. There would need to be very strong legal backing for forgetfulness tools to be mandatorily deployed, if they are to stand any chance of success in the world of Big Data.

This brings us full circle: legal measures to effect a right to be forgotten need to be backed up by technical measures if the right is to be enforceable in practice, while these technical measures need strong support from the law in order to be deployed.[92] Squaring this circle is not impossible, as we have seen with copyright law and Digital Rights Management systems with their backup legislation,[93] but the power relations and advocacy coalitions in privacy and data protection are distinctly different from those in the copyright field,[94] which considerably lowers the odds that policy-makers will adopt a strong law/technology-combo approach to effect the right to be forgotten.

In conclusion, we cannot be altogether optimistic when it comes to effecting a right to be forgotten in the form of a right to have data deleted in due time. The common vision in the literature and in European policy on the DPD review can be summed up well as follows:

---

[89] Registratiekamer and Information and Privacy Commissioner, *Privacy-enhancing technologies: the path to anonymity,* 9034632024 (v. 1), 9034632024 (v. 2) (1995).

[90] London Economics, *Study on the economic benefits of privacy-enhancing technologies (PETs)* (2010), at xvi.

[91] *Ibid*.

[92] In this respect, Mayer-Schönberger paints a far too rose-coloured picture when he claims that his proposal for ubiquitous expiry dates is modest because "in many instances [it] would require limited technical modifications" and can be done "without relying on new rights or new institutions" (V Mayer-Schönberger, see note 7 above, at 189).

[93] See art. 6 *Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society*, OJ [2001] L167/10.

[94] Cf the analyses of copyright and privacy in EJ Dommering and L Asscher (eds), *Coding Regulation. Essays on the Normative Role of Information Technology* (2006).

> We envision this right [to delete] as a combination of technical tools, legal regulation, and social norms and market pressure that will work in combination with other laws and technologies to promote individual control of personal information.[95]

Although good starting points exist in law and technology to enable data subjects to have their data deleted in due time, the proponents of a right to be forgotten hardly substantiate how this vision could effectively materialise. A closer analysis of the legal and technical measures reveals that a strong legal and policy effort is required; both to actually shape a true right to be forgotten for data subjects and to have these rights effectively enforced through mandatory forgetfulness-by-design, while enabling technologies to get a grip on data shadows are still in an embryonic stage.

Indeed, a discrepancy can be discerned in the literature between the analysis of the problem (basically, the fact that users have no control over personal data in the world of Big Data) and the envisioned solution (basically, to give users the rights and means to control their personal data in the world of Big Data). This solution will not work by simply offering rights and tools to users, if the underlying mechanisms of the problem are not simultaneously addressed. This calls for a much more careful analysis of the mechanisms underlying Big Data than is currently on offer. Big Data develops in a climate of mutually reinforcing social, economic, psychological, policy, and technological mechanisms that lead to the multiplication of data rather than towards limitation of data,[96] to a "fetishization of recording for recording's sake".[97] A plea for "individual control of personal information" is a welcome contribution to try and counter the trend of data proliferation, but sound arguments underlining the importance of forgetfulness and informational self-determination will not as such be able to effect a right to be forgotten. It remains to be seen whether there is indeed sufficient policy urgency in Europe to change data-protection law and to mandate forgetfulness-by-design to make a right to be forgotten as a right to have data deleted in any way meaningful.

4.2. A Claim on a Clean Slate

Although the right to be forgotten is usually cast a right to delete data in due time, two less dominant perspectives exist that focus on a "clean slate".[98] The first of these stresses the claim individuals may have that outdated negative information is not used against them. This is not only in the interest of individuals, but also of society: since people can change and circumstances also change, some areas of social life function better if people are given the chance to start from scratch. Several areas of law already have mechanisms to enable "social forgetfulness", such as bankruptcy, purging crime records and rehabilitation measures in juvenile criminal justice, and limiting the period during which negative facts can influence people's credit history in credit

---

[95] C Conley, see note 5 above, at 58.

[96] BJ Koops and R Leenes, "'Code' and the Slow Erosion of Privacy" (2005) 12 *Michigan Telecommunications & Technology Law Review* 115-188; BJ Koops, "Technology and the Crime Society: Rethinking Legal Protection" (2009) 1 *Law, Innovation & Technology* 93-124.

[97] LJ Bannon, see note 21 above, at 10.

[98] See s 2.2 above.

reporting.[99] Extrapolating the "clean slate" interest embedded in these laws to a world of increasing data retention, Blanchette and Johnson call upon policy-makers and organisations to be more aware of the social value of forgetfulness when setting legal, bureaucratic, or self-regulatory retention periods for all kinds of data.[100]

In a similar vein although in a different context, Werro casts the right to be forgotten in the form of the interest of criminal offenders not to be permanently confronted with their former crimes in the media; the trade-off between freedom of the press and the right to privacy should gradually put more emphasis on the latter as time passes. The right to be forgotten in that sense is a claim individuals have on the media not to publish negative facts about their past.

In this guise, a right to be forgotten builds less on concrete data-protection rights to have data deleted, than on the more abstract right to privacy, particularly in the form of "the freedom from unreasonable constraints on the construction of one's identity".[101] People must be able to shape their own lives, and therefore should not be fixed in the perception of others by their past.

In this form, the right to be forgotten can be invoked by individuals against parties who process data about their past, in particular who intend to publish these data or make decisions about people based on these data. It is typically an ex post right, in that it will usually be invoked after the publication or decision, to claim redress (although the shadow of this may of course have a preventative effect on the press or decision-makers). (It might also be invoked, perhaps, by individuals against the state when long legislative retention periods violate their right to privacy, although as noted in the previous section this only seems feasible in extreme cases.[102] For curbing legislative retention periods, forgetfulness seems more a policy guideline, as an important value to be taken into account by the legislature, than a subjective right.[103])

For effecting the right to be forgotten in this guise, it may not be necessary to change the law, as the interest of forgetfulness is already embedded in the general right to privacy as well as in several sector-specific legal provisions protecting vulnerable people from suffering unduly from past incidents. Nevertheless, the inexorable memory of digital footprints and data shadows seems to give rise to more people being harmed in different settings than those in which clean-slate protection currently exists. The right to be forgotten could therefore be extended from current areas such as bankruptcy law and juvenile justice to other areas in which people can suffer from being fixed in their past rather than judged on present merits. For example, labour law,[104] consumer law, and administrative and preventative criminal justice could

---

[99] See for a discussion of these: J-F Blanchette and DG Johnson, see note 6 above.

[100] *Ibid*.

[101] PE Agre, "Introduction" in PE Agre and M Rotenberg (eds), *Technology and Privacy: The New Landscape* (Cambridge, MA /London: The MIT Press, 1998) 1-28, at 7.

[102] See notes 56-57 above and surrounding text.

[103] J-F Blanchette and DG Johnson, see note 6 above.

[104] An interesting example is the German Labour Data Protection Bill [Entwurf eines Gesetzes zure Regelung des Beschäftigtendatenschutzes], which provides in §32(6) of the Federal Data Protection Act that in recruitment, employers are allowed to acquire publicly available data, but not data on social networking sites (e.g. Facebook) except sites that focus on professional qualifications (e.g. LinkedIn).

benefit from imposing limitations to how strong parties can use data to make decisions about weak parties.[105]

These could be in the form of measures limiting how long potentially detrimental data can be kept about individuals, but also in another shape. In some respects, the right to be forgotten can also be compared to the exclusionary rule in criminal evidence: some data have to be left out in judicial decision-making because they do not contribute to a fair decision. Just as the exclusionary rule forces courts to forget about evidence they have seen, the right to be forgotten could be constructed as rules forcing decision-makers to forget about data that are unduly detrimental to the individual they are affecting with their decision. For example, if Alice applies for a job, she will want her prospective employer Eric to disregard Bob's blog post and the incriminating photo tagged by Carol. An exclusionary rule, embedded in codes of conduct or in labour law, could safeguard that Eric cannot deny her the job on the basis of online information that is irrelevant for the job. Of course, Eric can say that Annie was simply more suitable for the job than Alice, but this situation is nothing new in labour law: an exclusionary rule, much like non-discrimination requirements, functions largely as a duty to motivate the decision, which allows the subject to challenge the decision if she feels it was unduly influenced by the information that should have been disregarded. Thus, the right to be forgotten could also be translated into sector-specific and context-specific norms specifying which online information should be included or excluded in decision-making. This could then be enforced through existing oversight mechanisms in the sectors at issue.

In conclusion, the right to be forgotten as a claim on a clean slate is much more modest and narrower than the right to delete that currently dominates the debate. It does not focus on comprehensive measures aimed at individuals being able to control which information exists, but rather on fine-grained, context-specific measures aimed at controlling how other parties can use information when decisions are made that affect individuals. It can be effected partly through existing legal rights – the right to privacy and some sector-specific rights to start with a clean slate – but may have to be extended to cover more areas in which people are particularly vulnerable to being unduly confronted with detrimental information about their past. This could be done through limiting periods during which detrimental data can be retained, but also through legal mechanisms similar to the exclusionary rule and non-discrimination oversight that enhance fair decision-making.

4.3. Unrestrained Individual Expression Here and Now

The other alternative, "clean-slate" perspective on the right to be forgotten is more philosophical and psychological in character, and stresses perhaps the right to forget rather than the right to be forgotten. Exemplified by people suffering from hyperthymesia (superior autobiographical memory), who "feel shackled by their constantly present past, so much so that it constrains their daily lives, limits their decision-making ability, as well as their capacity to forge close ties with those who

---

See Drucksache 17/4230, 15 Dec 2010, http://www.aus-portal.de/media/1704230.pdf (accessed 1 Nov 2011).

[105] Cf BJ Koops, "Law, Technology, and Shifting Power Relations" (2010) 25 *Berkeley Technology Law Journal* 973-1035.

remember less",[106] this vision stresses the importance of the here and now for individuals' experiences and social relationships. Rouvroy conceives of the right to forgetfulness largely as a right to speak and write freely, without fear of your person(ality) being fixed by what you express; it implies the sense of liberty of writing today and being able to change your mind tomorrow.[107]

This guise of the right to be forgotten resembles the claim-on-a-clean-slate perspective in that it aims at preventing people from suffering unduly from information about their past, with connections to the right to privacy and identity construction. However, it seems to have to function ex nunc (when data are created) rather than ex post (when data are used in decision-making). This makes it more indeterminate as to the persons against whom the right could be invoked. If you feel unduly restricted to do something in the here and now, out of fear that the digital footprints or data shadows resulting from your actions will fix your personality in the future, it is difficult to determine who to turn to for invoking your right. It could be a future employer, friend, insurance company, lover, police department, or neighbour that may confront you in future with the memory of your behaviour today, but of course you cannot – ex nunc – go to court to request all potential future data users to refrain from using, now or in an indeterminate future, the data generated by your current actions. The only feasible way of enforcing a right to be forgotten in this guise would be to go to court ex post, i.e., at the point when one is unduly confronted with data from the past; and perhaps this could have a general preventative effect on data processors, so that you and people like you can behave more freely without the shadow of future data misuse looming over your behaviour. But conceived thus, the right to be forgotten is no different from the (ex post) claim on a clean slate discussed in the previous section, and it has little leverage to be invoked in the here and now that is the key concern articulated by Rouvroy.

This implies that this third variant of the right to be forgotten does not have the character of a legal *right*, but rather of an *interest* or value. It serves as a philosophical and socio-psychological reflection on how we shape our lives in a world of Big Data. Such a reflection can feed into law and policy, but it does not aim as such to be cast as a legal right that individuals could invoke (other, perhaps, than as part of existing rights to privacy and freedom of expression). This does not make the perspective of unrestrained individual expression in the here and now any less valuable. The interest in being able to forget and to be forgotten is a lesson worth remembering for legislatures and policy-makers when they make decisions about large-scale data-collection and data-processing infrastructures and data retention periods. It is also an important lesson for designers, as Bannon argues. Rather than designing systems, such as architectures for life-logging, ubiquitous computing, and ambient intelligence, in such a way that they capitalise on technology's capacity for sheer unlimited data generation and storage, system designers would do well to realise the social function of forgetting, in particular of the human ability to "celebrate the fleeting moment".[108] Through "a critical perspective on the ever-increasing computerization of daily life", designers can realise that different design options are available, including those that

---

[106] V Mayer-Schönberger, see note 7 above, at 13.

[107] A Rouvroy, see note 8 above, at 25-26.

[108] LJ Bannon, see note 21 above, at 12.

respect the "being-here-now" rather than bringing in the past or preserving for the future.[109]

Altogether, however useful this reflection on the value of forgetfulness is, it does not seem to constitute a right to be forgotten, or to forgetfulness, in the legal sense. It seems very relevant for policy-makers and designers to take heed of, but is less relevant in a legal debate on how to shape a right to be forgotten.

5. Conclusion

Having analysed the implications of a right to be forgotten in its various guises in a world of Big Data, what can we say about the question what a right to be forgotten could or should look like? From the literature, I have identified three possible conceptualisations of the right, a dominant perspective stressing that personal data should be deleted in due time, a social "clean-slate" perspective that outdated negative information should not be used against people, and an individual self-development perspective that people should feel unrestrained in expressing themselves in the here and now.

The third perspective has little to do with a legal right as such. It highlights the importance of being able to forget and of being able to act without fear that your current actions may haunt you for the rest of your life. In that sense, it is an important reminder for policy-makers and system designers that they should not blindly follow technological opportunities for having ever more data recorded, to the detriment of living and decision-making in the here and now. But although important as a lesson, this does not seem a fruitful conceptualisation of a legal right to be forgotten as a separate entity.

Basically, then, we have two visions of shaping a right to be forgotten: a right to have data deleted in due time and a right to a clean slate. The first comes down to "a right to delete that gives individuals the ability to control their own history and thus escape it."[110] This vision is in line with the informational self-determination perspective that underlies current thinking about revising the Data Protection Directive.[111] It is a right that can be invoked against data controllers, possibly ex ante through users setting expiry dates after which data are automatically deleted, or ex post when users request deletion of certain data. This vision, however, is not as easy as it may sound. A closer analysis reveals practical and fundamental complications, partly due to the current limitations of data-protection rights (such as the household exception for many users uploading content about others in social networking contexts, and the limitation to erasure requests for inaccurate or unlawfully processed data),[112] but also partly due to intrinsic tensions in the right to delete.

One such tension is that ideally, the right has a prospective, preventative character, rather than only being invokable when data are doing concrete harm; however, it is

---

[109] *Ibid*; cf Dodge's and Kitchin's suggestions for designing in forgetting, see note 6 above and surrounding text.

[110] C Conley, see note 5 above, at 58.

[111] V Reding, see note 1 above, at 4 ("I believe we need to strengthen individuals' rights by ensuring that they enjoy a high level of protection and maintain control over their data").

[112] See more extensively s 4.1.3 above.

difficult for users to make considered choices on expiry dates, which do not work well in any way for data shadows (and third parties creating data shadows will not necessarily have the data subject's interest in mind when setting expiry dates), nor does it work well for data with statutory retention periods. Another tension is that the right is focused at having data deleted "in due time", but this is an intrinsically vague concept. What is "due" depends first on the perspective of interests: does it refer primarily to the legitimate processing, i.e., the data controller's interest in being allowed to retain the data while needed, or does it refer rather to the data subject's interest in having the data removed when they are, on balance, more detrimental than beneficiary for her to be processed? Second, if "due time" includes some element of the data controller's interest in legitimate processing, what is "due" will also depend on the definition of the purpose of data processing, which in the era of Big Data may tend both toward vague purpose-definition to allow for future uses and toward extending purposes through function-creeping mechanisms to secondary purposes. This seems the core of the tension inherent to a right to delete: do the legitimate interests of the data subject in desiring data to be removed override others' legitimate interests to process the data lawfully, i.e. on legitimate grounds for as long as needed to fulfil the primary purpose? That would be in line with the focus on informational self-determination, but it would deviate significantly from the current data-protection provisions, which have a more balanced approach to potential conflicts of interests between data subjects and data controllers.

Altogether, the vision of a user-controlled right to delete will require significant changes in law and in design-based enforcement, if the right to be forgotten is to be really effective for data subjects. A strong legal and policy effort will be required; both to actually shape a true right to be forgotten for data subjects and to have these rights effectively enforced through mandatory forgetfulness-by-design, while enabling technologies to get a grip on data shadows are still at an embryonic stage.

The second vision, involving a clean-slate perspective, perhaps suffers less from such complications and challenges, as it is more modest and narrow. This vision does not focus on comprehensive measures aimed at individuals being able to generically control which information exists, but rather on fine-grained, context-specific measures aimed at controlling how other parties can use information when making concrete decisions that affect individuals. It can be effected partly through existing legal rights – the right to privacy and some sector-specific rights to start with a clean slate – but because of the developments in Big Data, it may have to be extended to cover more areas in which people are particularly vulnerable to being unduly confronted with detrimental information about their past. For example, the right to be forgotten could be embedded not only in bankruptcy law and juvenile justice, but also in labour law, consumer law, and administrative and preventative criminal justice. This could be done in the form of limiting periods during which detrimental data can be retained, or through legal mechanisms similar to the exclusionary rule and non-discrimination oversight that enhance fair decision-making about job applicants and employees, consumers, (quasi-)suspects and administrative offenders. How exactly a right to be forgotten could or would need to be shaped in various areas of law needs to be developed further; since the "clean-slate" perspective is only a minority stream in current literature, the concrete shaping of a right to be forgotten in this sense is yet very underdeveloped.

The main differences between the approaches are summarised in the following table.

	right to delete in due time	**right to a clean slate**
object	deletion of data	blocking use of data
type	data subject right	data processor obligation
focus	data collection and storage	data use in decision-making
scope	generic	specific
legal area	data-protection law	sector-specific law
enforcement	enforcement-by-design	legal measures for oversight

Choosing between the two visions on shaping a right to be forgotten is a matter of outlook. I am sceptical of the first, comprehensive, user-control-based vision and tend to favour the narrower second, fine-tuned, clean-slate vision. Looking at the world of Big Data we live in, I tend to believe that the data-deluge genie is out of the bottle. No matter how important the ideal of informational self-determination may be, users will not be able to put it back again. I doubt whether there is sufficient policy urgency in Europe to substantially change data-protection law to give data subjects a full-blown right to have data deleted, and to simultaneously mandate the forgetfulness-by-design that is required to make a right to be forgotten in any way meaningful. However, scholars and policy-makers with a different outlook may feel differently, and aim for devising legal and technical solutions that can address the challenges I outlined for a user-controlled right to be forgotten.

In any case, it is clear that a generic right to be forgotten does not currently exist. There are flavours of such a right in current data protection and sectoral "clean-slate" laws, but the first are limited in strength, the second are limited in scope. Given the different possible conceptualisations and their different foci, anyone who advocates the establishment of a full-blown right to be forgotten must clarify what this right means and how it can be effected. As argued in this paper, considerable obstacles need to be overcome if people are really to be able to have their digital footprints forgotten and to shun their data shadows.