9 Squares: Framing Data Privacy Issues

Eerke Boiten

School of Computer Science and Informatics, De Montfort University, UK
(research carried out at University of Kent)

1. INTRODUCTION

Data protection has been a topic of intense discussion, particularly within the EU, for recent years. The agreement on the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679)[1] on 27 April 2016 was an important milestone in this, but some of the finer points in this are still up for discussion or left for implementation in the final text.

A particular point of contention is still the "*Right To Be Forgotten*", in both of its forms. This was proposed by European Commissioner Viviane Reding in 2010,[2] sparked off by Facebook's practice of not properly closing down user-deleted accounts.[3] It is defined as an extension of the principle of *data minimisation* which already existed in the 1995 Data Protection Directive[4]. The Right is introduced in the GDPR as follows[5]:

> *A data subject should have the right to have personal data concerning him or her rectified and a 'right to be forgotten' where the retention of such data infringes this Regulation or Union or Member State law to which the controller is subject. In particular, a data subject should have the right to have his or her personal data erased and no longer processed where the personal data are no longer necessary in relation to the purposes for which they are collected or otherwise processed, where a data subject has withdrawn his or her consent or objects to the processing of personal data concerning him or her, or where the processing of his or her personal data does not otherwise comply with this Regulation.*

The legal aspects of this new right have been debated extensively ever since the idea was first mooted, with concerns being expressed over this privacy right unduly impinging on freedom of speech or freedom of expression. Discussion intensified in May 2014 when the Court of Justice of the European Union gave their decision in the Costeja[6] ("Google Spain") case, which

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> (accessed 19 July 2016).

[2] V. Reding, 'Citizenship Privacy matters – Why the EU needs new personal data protection rules', The European Data Protection and Privacy Conference Brussels, 30 Nov 2010, <http://europa.eu/rapid/press-release_SPEECH-10-700_en.htm> (accessed 19 July 2016).

[3] M. Aspan, 'How Sticky Is Membership on Facebook? Just Try Breaking Free' (New York Times, 11 Feb 2008). <http://www.nytimes.com/2008/02/11/technology/11facebook.html> (accessed 19 July 2016, requires free registration).

[4] 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, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML>, Article 6 (accessed 19 July 2016).

[5] GDPR, recital (65).

[6] *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos (AEPD), Mario Costeja González*, <https://e-justice.europa.eu/ecli/ECLI:EU:C:2014:317> (accessed 19 July 2016).

led to “delisting” on the internet on the basis of the 2005 DPD, even before the GDPR was fully agreed on.

Because a balancing of different human rights is involved, the general answer to the question “Does the right to be forgotten apply to a specific piece of information?” is going to be “It depends”. The next question is then “On what?” UK researcher P. Bernal[7] unpicked this in relation to the different prevalent views between the US and the EU on this issue. He indicated different contexts types of data, including “stories” (as we use here) as well as referring to “data that is not really speech in any practical sense”. Google Global Privacy Counsel P. Fleischer[8] also noted that the lack of a precise framing of the issues muddled the discussion. This paper takes such ideas further by presenting a particular taxonomy of personal information and online resources, which helps frame discussions on the Right To Be Forgotten as well as other data privacy issues.

The rest of this paper is set out as follows. The next section introduces the taxonomy, by presenting the two dimensions and each of the three possibilities in each dimension. The section that follows discuss each of the resulting nine combinations separately, highlighting the issues that arise in each “square”. The paper ends with short concluding comments.

2. A TAXONOMY OF PERSONAL INFORMATION AND ONLINE RESOURCES.

For personal data in databases or on online resources (websites, social networks, etc.), the strategies and issues around privacy and in particular whether and how such information can be “forgotten” vary, depending on the nature of the online resource and the type of personal information concerned. We propose a simple two-dimensional analysis, leading to nine different scenarios.

2.1 First dimension: who owns the resource?

The first dimension characterises the type of online *resource* by how it relates to the individual. The three types we distinguish are as follows.

- **ME**: the individual is not just the *subject* but also the owner of the resource. A typical example of this is the pre-social network concept of a “personal webpage”, hosted on private web space or on some generic service (such as GeoCities) that does not prescribe the nature or detailed format of the information provided, nor does it expect any linkage between the contents of *different* individuals. Blogs and the contents of emails can also fall into this category.
- **US**: the individual is a *participant* in the resource, which presents itself as a kind of community. The data subject has probably agreed to terms and conditions in order to receive a service that is being provided. There is likely to be personal information provided confidentially to the resource (for identification, authentication, billing, etc.), as well as personal information provided through the resource to other participants, specific third parties, or the public. Some resources may not make any personal information available to the public, e.g. basic e-commerce websites. This reduces the relevance of such resources for discussions on privacy. However, note

[7] P. Bernal, ‘The EU, the US, and the Right to be Forgotten’, in: Gutwirth, S., Leenes, R.E., De Hert, P. (eds.), *Computers, privacy and data protection – reloading data protection*, Dordrecht etc.: Springer 2014.

[8] P. Fleischer, “Foggy thinking about the Right to Oblivion”, (9 Mar 2011) <http://peterfleischer.blogspot.co.uk/2011/03/foggy-thinking-about-right-to-oblivion.html> (accessed 19 July 2016).

that most more complex internet services tend to include aspects of "socialization" or reputation management that expose some personal information to the outside world. The structure in which the information is provided is mostly determined by the resource owner. The prototypical example of this is a social network, or the "comments" section of a website. There may be a disparity, however, between the cooperative way the resource presents itself to its participants and the reality of its business model.

- **THEM**: This is a type of resource that carries personal information about the individual where no relationship between them is established in advance; the individual is merely an "*object*" of interest. A typical example of this would be the "contents" section of a newspaper website, with stories about individuals that are considered to be of public interest.

2.2 Second dimension: what kind of data?

The privacy considerations also change depending on what kind of personal *data* is involved. We distinguish three categories of personal data:

- **ATT**ributes: this is the kind of data that is traditionally stored in databases. Items like name, date of birth, address, names of children – each of them occurring once or a low bounded number of times per individual, and often existing in a unique canonical abstract form. Pictures do not belong in this category, although passport pictures (by removing degrees of freedom or through their biometric measurements) come closer to this. Information given for administrative reasons to "US" resources is typically in this category. On a social network like Facebook, this kind of information when published is typically in a "Profile" or "About" section.
- "**STO**ries"[9]: this is all other *explicitly* generated personal information, such as medical history, pictures and other media, status updates on social networks, comments and posts on blogs and websites, and the contents of emails. Users can ask Facebook to provide "the full set of information" that Facebook holds about them[10] – the information returned will include these first two categories only, but exclude the final one.
- **BEH**aviour: this is all the *implicitly* generated personal information, such as location history as kept by smart phones and networks, metadata about email communication and web browsing, or purchase history information as collected by store loyalty cards. Facebook, for example, maintains its own search history, which is evident from typing a letter or two into its search window – but this information is not returned to customers on request, nor can it be reset or directly controlled in any way by users.

There is no doubt that all these three categories can contain *personal* data as in a Data Protection context. For example, location info has been shown to be highly useful for identifying

[9] P. Bernal, 'The EU, the US, and the Right to be Forgotten'.

[10] Europe versus Facebook, 'Get Your Data! Make an Access Request at Facebook!', <http://europe-v-facebook.org/EN/Get_your_Data_/get_your_data_.html> (accessed 19 July 2016).

individuals[11]: for identifying 95% of people, 4 data points were shown to suffice; the combination of work location and home location (up to block level) usually identifies uniquely[12].

The boundaries between the three categories may not always be sharp, and the same type of information can occur in each category: for example a location may be an attribute if it is someone's home address; a story if it is a check-in on Facebook; and behaviour if it is quietly recorded by a smart phone.

3. DATA PRIVACY: 9 SCENARIOS

The three types of online resources and the three kinds of data together lead to nine different combinations, which are summarised in the table below and discussed in some detail after that, proceeding column by column, and identifying the entries by the bold letters representing row and column.

data ↓	resource →	**ME** *subject*	**US** *participant*	**THEM** *object*
ATTributes		control but authentication effect to other contexts	traditional DP context	traditional privacy: attributes in public domain, sensitive?
"**STO**ries"		full control	typical social network content etc.	freedom of expression; cyberbullying
BEHaviour		some leakage (e.g. dynamics)	metadata, browsing, hidden data, surveillance capitalism	surveillance capitalism [13], data mining, deanonymisation

Firstly, in the **ME** column it looks like the data subject has full control over each of the types of data. However, even here we cannot currently claim that they are able to exercise a full right to be forgotten (or: erasure[14]) due to historical web archives such as the WayBack Machine[15] or more short-term archiving such as in Google caches. It *is* possible to remove and withhold

[11] Y.-A. de Montjoye, C.A. Hidalgo, M. Verleysen and V.D. Blondel, 'Unique in the Crowd: The privacy bounds of human mobility' [March 2013], Scientific Reports **3**:1376, doi 10.1038/srep01376. <http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html> (accessed 19 July 2016).

[12] P. Golle and K. Partridge, 'On the Anonymity of Home/Work Location Pairs', in H. Tokuda, M. Beigl, A. Friday, A. J. Bernheim Brush, Y. Tobe (eds): *Pervasive Computing, 7th International Conference*, LNCS 5538, pp 390-397, Springer 2009. doi 10.1007/978-3-642-01516-8_26.

[13] S. Zuboff, 'Big Other: Surveillance Capitalism and the Prospects of an Information Civilization' [2015], Journal of Information Technology 30, 75–89. doi:10.1057/jit.2015.5.

[14] P. Druschel, M. Backes, and R. Tirtea. 'The right to be forgotten – between expectations and practice' (European Network and Information Security Agency, November 2012), <http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/the-right-to-be-forgotten> (accessed 19 July 2016), section 3.3.

[15] 'WayBack Machine' <web.archive.org> (accessed 19 July 2016).

information from search engines, e.g. by asking Google for removal using special request forms, and using meta-tags to prevent indexing[16].

In terms of possible privacy invasions e.g. through Google search or "doxing" (the malevolent publication of personal data), the user's control over information in this column also has an impact on risks arising from the other columns. If, for example, a frequently occurring name creates ambiguity about whether other resources refer to this data subject, voluntarily published information in this column may serve to disambiguate that. Hence it not only authenticates the resource itself, but also external information. Such disambiguating information could include photographs, educational history and other information which reveals aspects of location or age, and family relationships.

ME-ATT: Although the individual has full discretion about which attributes to publish here, fully hiding them defeats the purpose of having (something like) a personal website. Displaying at least a name here serves as a kind of claim to authenticity and identification that this resource does indeed concern the data subject. Using a pseudonym might be an alternative, particularly in the context of one individual maintaining multiple online identities.

ME-STO: The stories are likely the "content" that is being published to the world in this scenario, and thus the raison d'être for the resource. With the proviso made in general for this column, the data subject is fully able to ensure that the current contents of the resource reflects positively on them, even if historical content provision might have been less well judged.

ME-BEH: There is some behavioural information leaking in this scenario, from the changes over time of what content is being provided. Mostly this will be inconsequential, and not worthwhile for third parties to track, but the individual cannot exercise control over this. Leaked information may be similar in nature to Facebook's helpful notifications concerning change in relationship status, e.g. – which in our experience have always reflected significant change of personal circumstances of the data subject.

Another context in which information ostensibly controlled by the data subject comes at risk is when hardware etc. is decommissioned or deleted. Securely wiping storage media on discarded computers is necessary to avoid leakage of information from all these categories of data.

In the **US** column, the data subject exercises their control over personal data in a way that is constrained and closely directed by the resource owner's procedures and intentions. The agreed terms and conditions or legal requirements in this context may also clarify which of the information provided by the subject may be shared with third parties or published.

US-ATT: This is the traditional scenario for the application of data protection legislation: a service that makes use of a database with personal information. There is a clearly defined and explicitly agreed relationship between data subject and data controller. The data subject knows what data is being requested, and can insist on data minimisation: that data is relevant for the purpose. The data subject may even be able to withdraw all their data from the resource by cancelling the service, unless the relationship is a legally required one, e.g. with the tax office. This is not an area where a right to erasure would add significant "consumer power".

US-STO: This describes the typical social network content, and is a main area of concern that The Right To Be Forgotten was aimed to address. As in the rest of this column, individuals may withdraw their data by cancelling the relationship with the resource owners, although the

[16] 'How to Ungoogle yourself', <http://www.wikihow.com/Ungoogle-Yourself> (accessed 19 July 2016).

latter may be reluctant to allow this[17]. Visibility of personal data to other participants, to the public including indirectly through search engines, and to third party businesses is controlled through terms and conditions and "privacy controls". Due to the large variety of data and complex structure of social networks, it is notoriously difficult to provide privacy controls which are both effective and user-friendly. Other types of online information for this category include "newsgroups" (the main discussion mechanism on the internet before WWW), which was subject to implicit and explicit social contracts as well as technological constraints. In particular, expiration and cancellation of users' posts could not fully be guaranteed due to the distributed (broadcast) nature of the medium. Data generated through personal health and fitness devices mostly fits into this category too: users are aware of (at least an abstract view of) the data that is produced, and may be able to publish or share this with other users of the platform. However, when such data has exploitable information that the users have no access to, it shifts into the next square. Again, straddling a boundary highlights a potential issue.

US-BEH: Many services in the **US** column are free to use for their participants, but this entry explains why they are nevertheless successful businesses. Data in this entry consists of browsing behaviour, metadata of all types of communications, etc., including in processed forms, e.g. aggregations and derived profiles which can be used for targeted advertising[18]. In some cases, the service attempts to decouple the information from the subject's identity, for example through pseudonymisation. There are many reports of successful de-anonymisation of such data sets[19] [20]. Although the General Data Protection Regulation makes it clear that pseudonymised data remains personal[21], it still recommends pseudonymisation as a protection for personal data.

Typical social networks provide very little privacy control over this data, nor do they make it explicitly visible to users on request. E-commerce sites give varying degrees of control and visibility on how past searching and purchasing behavior is used, for example in order to generate recommendations.

There is significant flow between this category and **THEM-BEH**: e.g. social networks both sell data to data brokers, and buy in more data from data brokers about their existing customers[22]

The **THEM** column represents a more problematic aspect of traditional data protection legislation: if there is no explicit connection between the data subject and the data owner, how can the data subject exercise their rights? They might not know in the first place that the data is

[17] M. Aspan, 'How Sticky Is Membership on Facebook? Just Try Breaking Free'.

[18] F.J. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Wolters Kluwer, 2015).

[19] A. Narayanan, V. Shmatikov. 'Robust de-anonymization of large sparse datasets.' [2008] IEEE Symposium on Security and Privacy, IEEE, 111-125. doi: 10.1109/SP.2008.33.

[20] M. Barbaro, T. Zeller Jr, 'A Face Is Exposed for AOL Searcher No. 4417749' (The New York Times, 9 Aug 2006). <http://select.nytimes.com/gst/abstract.html?res=F10612FC345B0C7> (accessed 19 July 2016, requires free registration).

[21] GDPR, Articles 6.4(e), 32.1(e), 89.1; recital (26): "Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person."

[22] J. Angwin, T. Parris Jr, S. Mattu 'Facebook doesn't tell users everything it really knows about them' (ProPublica, 27 December 2016) <http://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them> (accessed 17 January 2017).

being processed. One might consider data brokers, telemarketing companies, credit rating agencies in this category – or identity thieves and debt collecting firms; but also newspapers.

THEM-AT: In this area we find some classic privacy issues, as well as a dual security impact. When an individual's attribute information is not only held, but also ends up being published, e.g. a celebrity's info in a newspaper, was this assumed to be in the public domain; was it sensitive personal information? Is it in the public interest for e.g. a person's sexual orientation to be published? This area thus involves classic privacy dilemmas, including the tension with freedom of expression which comes into full force in the next category. "Doxing" also comes into this category. This may have an impact on physical security (publication of home address) as well as information security (date of birth, mother's maiden name).

THEM-STO: In this scenario, stories are recorded or published about an individual without their consent or involvement. Troublesome examples of this include defamatory false news[23] and cyber-bullying. In the latter context, we can look at Facebook's encouragement of "tagging" (i.e., mentioning of another participant in a way that links to their other information) as serving multiple goals. It not only increases connectedness of the "community", but it also moves some cyber-bullying from the THEM column to the US column. It gives the tagged person some control over being mentioned, as well as making mentions more structurally visible and in that way, cyber-bullying potentially easier to police. Hence it is not surprising that names getting tagged is implemented by Facebook as a quite coercive default with tags automatically suggested when typing. On Twitter, the phenomenon of "sub-tweeting" (reference to another account without using the available linking mechanisms) lives in the same sphere.
Looking beyond social networks, this is where the Right To Be Forgotten is traditionally seen to clash with freedom of expression. The press, blogging individuals, and factual websites want to publish stories about other people; giving the subjects of these stories the right to have them withdrawn would be a clear attack on freedom of expression, and in some cases "allowing the re-writing of history". It is not clear to us what a right to be forgotten as applied to *published* information could sensibly achieve in this context beyond what is already attainable through existing limits on freedom of expression such as libel laws, child abuse prevention, or copyright legislation. Or, for that matter, data protection constraints on publication where the public interest is not seen as significant enough to judge it "fair processing". However, a right to deletion is very relevant for the situations where third parties hold personal data for other purposes than publication. One might consider any indiscriminate "hoovering" of email content by NSA/GCHQ's (as well as Google's) to fit within this category, for example …

THEM-BEH: The gathering of behavioural personal data by third parties is really the "1984" scenario of a surveillance society, of which there are ongoing and abundant examples: on a small scale, there were the waste bins in London which collected data of passing smart phones[24], and more seriously, the mass collection of communications metadata through surveillance by a variety of government-related organisations as legislated for in the UK and elsewhere. The company doing the waste bin phone tracking claimed they were only recording anonymous data (through MAC monitoring). Data protection legislation applies in this area,

[23] P. Oltermann. 'Syrian who took Merkel selfie sues Facebook over 'defamatory' posts' (Guardian, 12 January 2017), <http://www.theguardian.com/world/2017/jan/12/syrian-who-took-merkel-selfie-sues-facebook-over-defamatory-posts> (accessed 17 January 2017).

[24] 'City of London calls halt to smartphone tracking bins' (BBC News, 12 August 2013), <http://www.bbc.co.uk/news/technology-23665490> (accessed 19 July 2016).

but the lack of transparency and the scale of processing[25] make monitoring and enforcement extremely hard. Data brokers are at the heart of this category – starting by acquiring data from the US column, then combining several sources of data, and applying data mining techniques on them. Even if some of these initial data sets could have been reasonably viewed as "anonymized", profiling and data mining may effectively lead to the generation of personal data. The GDPR provides clarity and privacy in this area in two ways: by branding pseudonymisation as a security control rather than a shift out of personal data[26]; and by explicitly identifying profiling as a relevant data processing activity[27]. The lack of control that data protection legislation, individuals, and society at large have over this "surveillance economy" is one of the largest problems of our modern information society. Increased availability of person-related data, for example through the internet of things, will only make this worse.

4. CONCLUDING COMMENTS

We believe that the examples given in the several categories, and the different flavours of the issues arising, indicate that the taxonomy described here can be used to meaningfully frame data privacy discussions. The taxonomy has been used in practice to explain data privacy issues to undergraduate and postgraduate taught students.

It is worth considering whether the boundaries between the various areas defined here are indeed sharp and meaningful. In many of our examples, a particular data situation which apparently straddles a boundary also embodies an issue worth of discussion. However, the definition of the boundary between US and THEM deserves some further consideration. If the distinction is whether there is an explicit relationship between the data holder and the data subject, then that ties in with particular underpinnings for fair data processing – in particular, consent and contracts; although that does not mean that all data processing in the THEM column is illegal under the GDPR. However, by including processing justified by legal obligations, e.g. by government organisations, in the *US* column, we appear to be focusing on the data subject's *awareness* of the data processing in drawing the boundary rather than on the legal justification. This also stretches the explanation of the explicit relationship between data subject and data holder: in such a situation "terms and conditions" are at best implicit. Where social networks are acting as data brokers and advertising networks, their position can be argued to be in both US and THEM, as the data subject's presence and control over their data fades in that scenario. The data crossing the boundary between US and THEM, within a single platform, or as a consequence of data brokering or "data sharing", presents an appealing metaphor. Where the GDPR insists that the right to erasure extends to third parties where they process personal information that had been made public by the data processor[28], that can be viewed as an attempt to widen the US area.

Finally, can this taxonomy be refined by introducing additional dimensions? From the discussions of the various areas, an obvious candidate for this is whether personal data is merely held, or also being made public. One area where the GDPR reflects a changed world since the DPD,

[25] F. Pasquale, *The Black Box Society – The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015).

[26] GDPR, recitals (26) and (28): "The application of pseudonymisation can reduce the risks ...".

[27] GDPR, recital (24): "profiling a natural person, particularly in order to take decisions…".

[28] GDPR, Article 17.2.

is in its acknowledging internet publication of personal data[29], even if the legal consequences of this are still limited. In our taxonomy, this distinction pops up particularly in the US column where social networks are seen to hold ATTR type data for administrative purposes as well as for publication, and both are treated significantly differently. For data that is not published, the distinction between STO and BEH data becomes a little more vague: stories are generated in the first place for publication. How, for example, should we classify Facebook's search history of their participants? (For the moment, the answer is probably "both" – as from its auto-completion suggestions, it is clear that Facebook keeps its own copy even after the user has explicitly deleted their search history.)

ACKNOWLEDGEMENT

I would like to thank the TRILCON 2016 audience and the students of CO643 at the University of Kent for their comments and questions. I appreciate the many useful questions and comments of the reviewer.

REFERENCES

1. J. Angwin, T. Parris Jr, S. Mattu 'Facebook Doesn't Tell Users Everything It Really Knows About Them' (ProPublica, 27 December 2016) <http://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them> (accessed 17 January 2017).

2. M. Aspan, 'How Sticky Is Membership on Facebook? Just Try Breaking Free' (New York Times, 11 February 2008). <http://www.nytimes.com/2008/02/11/technology/11facebook.html> (accessed 19 July 2016, requires free registration).

3. M. Barbaro, T. Zeller Jr, 'A Face Is Exposed for AOL Searcher No. 4417749' (New York Times, 9 August 2006). <http://select.nytimes.com/gst/abstract.html?res=F10612FC345B0C7> (accessed 19 July 2016, requires free registration).

4. 'City of London calls halt to smartphone tracking bins' (BBC News, 12 August 2013), <http://www.bbc.co.uk/news/technology-23665490> (accessed 19 July 2016).

5. P. Bernal, 'The EU, the US, and the Right to be Forgotten', in: Gutwirth, S., Leenes, R.E., De Hert, P. (eds.), *Computers, privacy and data protection – reloading data protection* (Dordrecht etc., Springer 2014).

6. P. Druschel, M. Backes, and R. Tirtea. 'The right to be forgotten – between expectations and practice'. European Network and Information Security Agency, November 2012, <http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/the-right-to-be-forgotten> (accessed 19 July 2016).

7. 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, <http://eur-

[29] GDPR, recital (6): "Natural persons increasingly make personal information available publicly and globally".

lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML> (accessed 19 July 2016)

8. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> (accessed 19 July 2016)

9. Europe vs. Facebook <http://europe-v-facebook.org/>.

10. P. Fleischer, ‘Foggy thinking about the Right to Oblivion’ (9 March 2011) <http://peterfleischer.blogspot.co.uk/2011/03/foggy-thinking-about-right-to-oblivion.html> (accessed 19 July 2016).

11. P. Golle and K. Partridge, ‘On the Anonymity of Home/Work Location Pairs’, in H. Tokuda, M. Beigl, A. Friday, A. J. Bernheim Brush, Y. Tobe (eds): *Pervasive Computing, 7th International Conference*, LNCS 5538, pp 390-397, Springer 2009. doi 10.1007/978-3-642-01516-8_26.

12. *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos (AEPD), Mario Costeja González,* <https://e-justice.europa.eu/ecli/ECLI:EU:C:2014:317> (accessed 19 July 2016).

13. Y.-A. de Montjoye, C.A. Hidalgo, M. Verleysen and V.D. Blondel, ‘Unique in the Crowd: The privacy bounds of human mobility’, Scientific Reports 3:1376, doi 10.1038/srep01376, March 2013. <http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html> (accessed 19 July 2016)

14. A. Narayanan, V. Shmatikov. ‘Robust de-anonymization of large sparse datasets.’ *IEEE Symposium on Security and Privacy*, IEEE, 111-125, 2008. doi: 10.1109/SP.2008.33

15. P. Oltermann. ‘Syrian who took Merkel selfie sues Facebook over 'defamatory' posts’ (Guardian, 12 January 2017), <http://www.theguardian.com/world/2017/jan/12/syrian-who-took-merkel-selfie-sues-facebook-over-defamatory-posts> (accessed 17 January 2017).

16. F. Pasquale, *The Black Box Society – The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015).

17. V. Reding, ‘Citizenship Privacy matters – Why the EU needs new personal data protection rules’, The European Data Protection and Privacy Conference Brussels, 30 Nov 2010, <http://europa.eu/rapid/press-release_SPEECH-10-700_en.htm> (accessed 19 July 2016).

18. ‘WayBack Machine’, <web.archive.org> (accessed 19 July 2016).

19. Wikihow, ‘How to Ungoogle yourself’, <http://www.wikihow.com/Ungoogle-Yourself> (accessed 19 July 2016)

20. S. Zuboff, ‘Big Other: Surveillance Capitalism and the Prospects of an Information Civilization’ [2015], Journal of Information Technology 30, 75–89. doi:10.1057/jit.2015.5.

21. F.J. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Wolters Kluwer, 2015).