EJLT European Journal of Law and Technology

A Right to Delete?

Paul A. Bernal [1]

Cite as: Bernal, P.A., 'A Right to Delete?', European Journal of Law and Technology, Vol. 2, No.2, 2011

ABSTRACT

This article looks at one of the potential for the establishment of a right to delete personal data. It asks whether a right to delete would be an appropriate, effective and proportionate tool in the context of personal data and if so what form should it take and what kind of an impact might it have.

One version of the idea of a right to delete, the 'right to be forgotten', is of particular current interest: its mooted inclusion in the forthcoming revision of the Data Protection Directive has produced much debate and comment, some of it extremely negative, some emotional and some displaying both ignorance and misunderstanding. This article will argue that the right to be forgotten needs to be renamed and recast in order to address these negative reactions and the real concerns that underlie them.

The article further argues that a qualified 'right to delete' should reflect a paradigm shift in attitudes to personal data on the internet: that the default should be that data can be deleted, and that those holding the data should need to justify why they hold it. This could help to shape a more privacy-friendly future for the internet, one that could provide a better balance between the needs to individuals for privacy, businesses for financial success and governments for security than currently exists.

1. Introduction

Personal data has proliferated on the internet in recent years - and as it has, so have the issues and problems that surround it. Accompanying those issues have been suggestions as to possible ways to address them, or at least how to reduce the problems associated with this proliferation of data. This article looks at one of these suggestions: the establishment of a right to delete personal data. The article asks whether a right to delete would be an appropriate, effective and proportionate tool in the context of personal data and if so what form it should be and what kind of an impact might it have. One version of the idea of a right to delete, the 'right to be forgotten', is of particular current interest: its mooted inclusion in the forthcoming revision of the Data Protection Directive (Directive 95/46/EC) has produced much debate and comment, some of it extremely negative, some emotional and some displaying both ignorance and misunderstanding. This article will argue that the right to be forgotten needs to be renamed and recast in order to address these negative reactions and the real concerns that underlie them. The article will further argue that a qualified 'right to delete' should reflect a paradigm shift in attitudes to personal

data on the internet: that the default should be that data can be deleted, and that those holding the data should need to justify why they hold it. This could help to shape a more privacy-friendly future for the internet, one that could provide a better balance between the needs to individuals for privacy, businesses for financial success and governments for security than currently exists.

1.1 The right to be forgotten - and reactions to it

The EC Communication of November 2010, which sets out 'a comprehensive approach on personal data protection in the European Union' refers to the right to be forgotten as:

> '... the right of individuals to have their data no longer processed and deleted when they are no longer needed for legitimate purposes.' [2]

The communication states that in the review they will examine ways to clarify this right: as shall be discussed below, set out in this way, the 'right to be forgotten' is not exactly a new right, but a right derived from the existing data protection principle of data minimisation. Despite this, the inclusion of the right within the Communication, together with its mention in speeches by Commissioner Viviane Reding, [3] particularly in the context of social networking services, produced a number of reactions in both media and political circles.

The idea of a 'right to be forgotten' has its origins in the French and Italian legal concept of a 'right to oblivion' - in French the 'droit à l'oubli', in Italian the 'diritto al' oblio' - which been described as 'the right to silence on past events in life that are no longer occurring' [4] such as crimes for which the person has later been exonerated. It has arisen through a combination of legislation and jurisprudence since the late 1970s. [5] In this form it could be seen as restricting free speech - controlling what can and cannot be said in a particular way, albeit in terms that refer to legally established facts and events. As noted above, however, the online version of the right to be forgotten as set out in the Communication, is neither intended to be like that nor could function like that, at least in the terms set out in the Communication. It deals with the deletion of data that is no longer needed, rather than anything as dramatic as the erasing of past events or preventing any kind of speech. Nonetheless, the reaction to its suggestion seemed to a great extent with the older concept rather than what was mooted, suggesting ideas of rewriting history, of censorship, of gagging journalists, of using power to restrict free speech. Tessa Mayes, for example, writing in The Guardian in March 2011, [6] equated the right to be forgotten with a desire to 'live outside society' and suggested that if enacted it would 'signify the emasculation of our power to act in the world'.

Whilst it is hard to disagree with Andrew Murray's characterisation of Tessa Mayes' article as 'stunningly under-researched and naïve' [7] that may be precisely the point. It indicates one of the primary problems with the concept of a right to be forgotten: that it provokes emotional and instinctive reactions, often very negative, rather than rational and thought-through responses. Mayes' article was in what is generally considered to be an intelligent and high profile newspaper, and her thoughts were taken seriously enough for her to be invited to take part in a panel discussion at the Westminster Media Forum on the subject. [8] What is more, it is not just journalists that have reactions in this kind of direction: this author has presented papers on the subject at three different academic conferences in the last two years, and each time there has been at least one reaction

against the concept along these lines, and at times from distinguished and well-informed members of the audience. [9]

Political reactions have been more measured - but in some ways similarly negative and similarly dealing with issues not actually included within the right as suggested by the Commission. Lord Chancellor and Secretary of State for Justice Kenneth Clarke, in a speech to the British Chamber of Commerce in Belgium in May 26th said that 'other voices than mine have raised concerns over its ability to impinge on free speech, and to censor information which has been legitimately circulated in the public domain', and went on to raise further concerns about it limiting the portability of health records and to restrict the availability of credit histories - neither of which would be impacted upon by the right as suggested by the Commission, as such data would still be needed for legitimate purposes. Clarke further suggested that the right would be effectively unworkable as a result of the ease with which data may be replicated and shared. [10] It is a little hard to imagine that Clarke was not aware that his concerns over health records and credit histories were highly unlikely to have any foundation in relation to the online version of the right: it seems more likely that he was tapping into the vein of worries of those such as Mayes in order to support his broader concerns with the possible expansion of data protection as a whole, particularly given the audience to whom he was speaking.

1.2 Emotional reactions matter

The often emotional reactions to the idea of a right to be forgotten may not seem immediately of great importance in relation to law - but particularly in this context emotional reactions do matter, for a number of reasons. First of all, they matter politically, because there are many political hurdles that need to be tackled before this kind of law could be enacted, and politicians are acutely aware of the need to work at the emotional as well as the rational level - and are far from averse from using emotional means to get their other objectives achieved, as Clarke's speech referred to above suggests. Secondly, they matter in the battle for the hearts and minds of ordinary people. Articles like Mayes' or speeches like Clarke's can shape public views - changing public perceptions of what is being proposed and creating a groundswell of opposition to something that is quite different from what is perceived. The public matters - particularly in the context of the regulation of the internet. As the theories of network communitarians such as Murray and Scott suggest, the online community plays an active and significant role in the way that the internet is regulated, often more effectively than laws or lawmakers. [11] Examples such as the fall of Phorm's 'Webwise' behavioural advertising system, principally as a result of pressure from privacy advocates and the online community despite support from both government and business, [12] and Facebook's abandonment of its 'Beacon' system of data sharing for advertising purposes in the face of similar levels of resistance, show quite how powerful the internet community can be when it is mobilised. [13]

One final way in which reactions need to be considered very seriously is the role of the United States, particularly in relation to free speech. The key players in the internet world, particularly in relation to personal data, are principally U.S. companies: Google, Facebook, Microsoft, Twitter and so forth. If the idea of a right to be forgotten is automatically or emotionally associated with restrictions on freedom of speech, then those companies are likely to oppose it - as free speech is close to sacrosanct in the U.S., as the primacy of the

First Amendment requires. That opposition can have a great effect - and conversely if those companies can be convinced to get behind a move towards greater privacy or user control, then that move can have a real effect throughout the world. Google in particular uses the same standards for their users throughout the world, and this can sometimes be to the benefit of users' privacy. When they reduced their retention periods for search data to 9 months after pressure from the Article 29 Working Party, they reduced those periods for all users worldwide, not only those resident in the European Union. [14] For Google and Facebook to be convinced to comply with or cooperate with a right to delete data, it would have to be seen as consistent with rather than in opposition to freedom of expression - and for any kind of right to delete to function effectively on the internet it would have to have the cooperation of Google and Facebook.

1.3 Underlying issues

All these concerns seem to suggest that the implementation of a 'right to be forgotten' would be fraught with problems. It could face resistance from the media, from politicians, from the big players of the internet - and potentially from any number of other businesses operating online. It is important to acknowledge that although these reactions are sometimes emotional and sometimes based on a misunderstanding of what is being proposed, they do reflect significant and relevant concerns. Fears of censorship, of rewriting history - and of losing more through the introduction of the right that might be gained are real fears, and must be understood and where appropriate addressed. However, even considering these concerns, there are real issues that the right to be forgotten is intended to address. The amount of personal data gathered and held on the internet is enormous. The existence of that data itself is of concern - and people can feel that their privacy is being invaded. What is more, that data appears to be increasingly vulnerable: vulnerable to misuse by those who gather it, vulnerable to acquisition by governments through legislation or court action, vulnerable to hackers or other criminals, vulnerable to those who might leak it for good reasons or bad, vulnerable to sale or other commercial misuse, vulnerable to function creep. It can be aggregated or combined with other forms of data for profiling.

The extent to which data can be legitimately used in ways that those about whom the data has been gathered would neither understand nor desire is wide-ranging. Even when data is gathered with legal consent, the data subjects will not always (or even often) understand that consent - often having scrolled through pages of legal language that they don't even read let alone understand before clicking 'OK'. Once this consent has been given, what happens to the data is effectively beyond the control of the user - it may be passed to third parties (within or without the terms of the consent), the use may shift (again, within or without the terms of the consent), and even the nature of the owners may change - for example as a business model evolves, or even if a company is taken-over by another company. Data can be taken from a company by governmental authorities through various legal means from subpoenas to the use of various forms of legislation. Data protection law can protect the data from some of these risks - but for a great many of them it is effectively powerless, partly as a result of the nature of consent as noted above, partly as a result of the difficulties that data protection authorities have in detection and enforcement.

Risks outside the law are also extensive - most dramatically through hacking. A particularly graphic example occurred in April 2011, when hackers attacked the Sony Playstation Network and stole the personal details of more than 100 million users. [15] These details include names, home addresses, email addresses, dates of birth, phone numbers, gender information and 'hashed passwords' - and in some cases direct debit details, credit card numbers and expiry dates. The direct debit and credit card details came from what Sony described as an 'outdated database' [16] - which in itself raises a lot of questions, most directly why that database even existed, let alone was accessible online for hackers. When the nature of Sony is considered, the hack is very revealing. Sony should be amongst the most technologically advanced and sophisticated organisations, with access to the best experts in security and in particular network security - and yet they were hacked, and hacked with great success. If Sony can be hacked, is anyone secure?

Sometimes it does not take particularly great technical expertise to access information on the internet. In May 2011, Matthijs R. Koot, a PhD student in the Netherlands, used simple techniques to mine Google's databases and put together a database of 35 million Google users including names, email addresses and biographical details. As Koot put it, it was 'completely trivial for a single individual to do this,' [17] and the process was completely within Google's rules, as they allow indexing of their public user information. Similar examples of all the wide variety of different kinds of vulnerability can be shown, from the HMRC child benefit data disc loss in 2007 [18] to the skimming of private WiFi networks by Google StreetView cars in 2009 [19], and the use of stolen data from Swiss banks by the German, French and UK governments to root out tax evaders. [20] Taking it further, over the last two years the activities of Wikileaks and hacker groups such as Anonymous and LulzSec [21] have in their very different ways emphasised the vulnerability of data - and how easily, once the data is compromised, it can be spread across the internet and across the world. Additional complications such as the rise of cloud computing make it even harder to keep data under control.

Does this all matter? It does appear to matter to people if evidence from the ICO is to be believed. In their 2010 'Response to the Ministry of Justice's Call for Evidence on the current data protection legislative framework' the ICO revealed that their research indicated that 'individuals increasingly feel they have lost control of their personal information'. [22] Ultimately, it is a question of autonomy. If people's most personal information can be so easily lost, and potentially put into the hands of criminals or others who could or would wish to use it against them, people feel in danger. If their data is vulnerable, the people themselves are vulnerable. If their data is threatened, people themselves feel threatened. The use (and misuse) of data can result in direct threats to autonomy - but it is perhaps equally important to understand that there is a feeling of a threat to autonomy that is of great importance too. If the problems are to be addressed, they must be addressed at both levels - people must both have more control over their data and they must feel that they have more control over their data.

2. What can be done?

If a right to be forgotten, as it is currently posited, is not a practical proposition - or at the very least faces enormous challenges if it is to have any chance of becoming a reality - what can be done to address these real issues? Before returning to the idea of a recast

'right to delete', other ways to address the issue should be considered. In particular, in what other ways can data be made more secure - and people made to feel that their data is less vulnerable, and hence that they themselves are less vulnerable.

2.1 Developing existing law and practice

The first and most direct way to deal with data vulnerability is through the development of existing law and practice. This could begin with the suggestions by the ICO of increased fines and harsher sentences to deal with data losses and failures of data security, and in particular, the possibility of custodial sentences. The possible penalties have recently been increased - from April 2010, fines could be as great as £500,000 [23] - but to date custodial sentences have not been introduced. It seems quite possible, however, that these kinds of penalties could eventually be brought in. The Information Commissioner also suggested that the possibility of extradition in appropriate cases should be opened up - given the nature of the internet that would again seem logical and appropriate. Whilst there are benefits to these ideas in terms of deterrence, there are also significant weaknesses. The question of whether deterrence really 'works' is not within the scope of this article but it is at least fair to say that, like the better use of encryption and other technological security measures which will be discussed below can only offer part of a solution to the problem. Recent experience - and in particular the 2011 ACS:Law case where although fines of £200,000 were initially threatened, only £1,000 was eventually levied as a result of the sole-trading solicitor involved winding his company down [24] - also suggests that even in the most direct cases, and even after stronger powers have been granted, the ICO may not apply them. So long as this is the case, the chance of deterrence is even less likely. Further to that, deterrence can only have a chance of functioning if the potential offenders believe that there is a significant likelihood of their being caught - and that is only likely if the enforcement arms of data protection authorities are substantially strengthened.

An improvement and strengthening of that enforcement is something that could potentially make a difference, improving data security and reducing data vulnerability. Even so, the nature of current law and practice, and the principle of proportionality mean that this kind of law - and in particular the idea of penalties harsh enough to act as a deterrent - could only apply to significant breaches and clearly 'sensitive' data. The problems relating to data vulnerability do not just apply to large scale events or to directly sensitive data - the vulnerability of seemingly innocuous data is also important, and the accumulation and aggregation of individually insignificant pieces of data can also have a significant impact, something that will be looked at in more depth in 3.4 below. These kinds of breaches are not only less likely to be detected but even if they are detected are highly unlikely to incur substantial penalties. More to the point, it would not be appropriate for them to do so. The problems with them arise through their accumulation rather than from each individual breach.

2.2 Better use of technology

There are technological tools that can help with data security - the most obvious being the use of encryption. A proper discussion of the use of encryption is beyond the scope of this article, but it is clear that encryption is a powerful tool in the practice of data security. However, it is also important to understand that the real experts do not believe that

encryption is anything more than a tool in the overall scheme of things. Ross Anderson, Professor of Computer Security Engineering at Cambridge University, and one of the leading experts in cryptography in particular and computer security in general, when asked 'How well-encrypted must data be, in order to be safe?' replied:

> 'You are in a state of sin. This is a wrong question to ask, for many reasons. `Whoever thinks his problem is solved by encryption, doesn't understand his problem and doesn't understand encryption' (Needham and Lampson)' [25]

What is more, even encrypted data is potentially vulnerable in two different ways. Firstly, the encryption itself can potentially be hacked or broken - there is an ongoing battle between the developers of encryption technology and the hackers trying to break it. Any code can and will eventually be broken - the question is whether those who are attempting to keep the data secure stay ahead of those who are attempting to break it. A further implication of this, and a further potential weakness, is that it requires those who use encryption to keep that encryption up to date - which leaves further scope for human error. That leads to the second weakness - that the use of encryption requires human interaction, and even if the encryption cannot be 'broken', sometimes the human can. As Ross Anderson puts it:

> 'As designers learn how to forestall the easier techie attacks, psychological manipulation of system users or operators becomes ever more attractive' [26]

Most directly, people might be persuaded either to release the keys to their encryption or even not to use the encryption properly at all. The use of psychological or emotional manipulation, particularly on the internet, is a developing science. This is Ross Anderson again:

> 'Deception, of various kinds, is now the greatest threat to online security. It can be used to get passwords, or to compromise confidential information or manipulate financial transactions directly.' [27]

So what does all of this mean? Simply that though technological tools are a crucial part of the process of improving data security and reducing data vulnerability they are far from being the whole solution.

2.3 Changes in the community and culture

An overriding requirement is that all the issues concerning information vulnerability and security be taken more seriously at every level. That must start from the very top. The ICO's position paper, 'Taking Stock, Taking Action', issued in the aftermath of the HMRC disc loss and the other data breaches that came to light following it, suggested that a 'role should be created at board level in larger organisations to deal specifically with information risk', and that '[a] post at senior executive level should oversee information security'. [28] The changes must be reflected throughout the organisation, and include proper and professional information risk management policies, periodic reporting of information risk at board level, clear lines of accountability and so forth, together with proper staff training and support. This is clearly of great importance, and a crucial first step towards an environment where data vulnerability can begin to be reduced.

The possibility of culture change can be taken to another level. The ICO position paper following all the reviews focused on that awareness simply in terms of the individuals' roles as employees of their organisations, but the real problem and indeed the potential solution runs deeper. If people were more aware of the issues of data security and vulnerability - and indeed of data privacy - in their ordinary personal and social lives, then it would be far easier for them to understand the importance of data security in their professional lives. They would find it easier to understand and implement information security policies, they would care more if the data encryption systems on their computers weren't functioning properly, and they would be less likely to fall for the kind of deceptive practices used by sophisticated cyber criminals. This culture change is perhaps the single most important factor - but it is also a factor that is very difficult to change, and something likely to take a considerable amount of time. There are signs that it may be happening, but at the same time there are suggestions of precisely the opposite - Facebook founder Mark Zuckerberg's suggestion at the Crunchie Awards that 'privacy was no longer a social norm' [29] is just one of many who have followed Scott McNealy's famous quote that 'You have zero privacy anyway. Get over it.' [30]

2.4 Taking data minimisation seriously

Even when these other factors are taken into account - if the law is both improved and better enforced, if the culture of organisations is more 'data-conscious' and where technology is used appropriately and effectively - there will still be problems, and risks that cannot ever be completely eliminated. Human errors, human nature, human malice, technological error and technological developments, and community pressures such as the demands to fight terrorism or catch child abusers or murderers are just some of the possibilities. Ultimately, wherever data exists, it is vulnerable - so the only way that data can really not be vulnerable is for it not to exist. Blogger Harry McCracken, when talking about the vulnerability of data held on Facebook said:

> 'Facebook has a history of asking for forgiveness rather than permission, and now says the default for everything is 'social'- so the best way to keep things private is to keep them off the service, period.' [31]

McCracken's argument can be extended not just to cover Facebook, but the whole of the Internet. The default for the whole of the Internet is that everything is 'public': the best way to keep things private is to keep them off the Internet completely. Taking it one step further, the best way to keep things private is not to keep them in a digital - and hence vulnerable - form. The ultimate weapon for in the fight against data vulnerability is to eliminate the very existence of data wherever possible.

The starting point for this is stronger, better-understood and better-implemented data minimisation. The concept of data minimisation is built in to data protection law. It combines the third and fifth data protection principles, as set out in Schedule 1 to the Data Protection Act 1998: [32] that data should be 'adequate, relevant and not excessive' and 'not kept for longer than is necessary'. It is, however, a concept that seems to be paid far less attention too than it should, partly, perhaps, because the terms are very difficult to define. What is 'excessive' and how long is 'necessary'? In specific cases the point has been argued at length by European regulators - for example in the Article 29 Working

Party opinion 148 concerning search engines [33] - but in general the answers to the questions are left to the discretion of those holding the data. Unless specifically challenged, the holders can choose how much data to hold and how long to hold it for - and as things stand, it appears that many businesses choose to hold more data than they need and for longer than they need to. What is more, data minimisation is scarcely enforced - and is in some ways inherently difficult for authorities to enforce. Authorities would have to institute some kind of compulsory 'data audit' in which they examine data policies and practices of anyone holding data - the difficulties and costs surrounding anything like this would make it all-but impossible. Moreover, the idea of imposing penalties for failures to appropriately minimise data would seem a step too far in the current data protection climate. As noted in 2.1 above, though the ICO now has the power to impose significant financial penalties, these can only apply for the most dramatic of data breaches - to extend this to penalties likely to have any effect on data holding policies in the round would be close to inconceivable.

The best way - perhaps the only way - for things to change positively in this field is for a new business model to develop. The key is to find a way to encourage the development of new business models that get closer to a real sense of data minimisation. How could this happen? If a way can be found to put the data subjects more in control of the data minimisation process, then not only will people be more in control of their own data but businesses would be put in a position where they have to develop these business models, business models that do not depend on their ability to gather whatever data they choose and hold it as long as they would like. That brings us back to the idea of a right to delete.

3. A paradigm shift in privacy

One of the principle aims of rights in general is to put power into the hands of individuals, power that can and should restrict the actions of those who might oppress, abuse or take advantage of those individuals. That kind of transfer of power, that kind of re-balancing, could have possibilities to redress the current imbalance over personal data - and to help to re-establish at least some of the control that people both have lost and feel that they have lost. Granting one group rights imposes duties on others. As noted at the start of this article, as the European Commission spells it out, though individuals do not currently have a 'right' to be forgotten, it can be argued that those holding the data do currently have a duty to 'forget' them. All that the right to be forgotten consists of, in the simple form as set out in the Communication, is 'the right of individuals to have their data no longer processed and deleted when they are no longer needed for legitimate purposes'. As noted above, data minimisation already requires those holding data not to hold it any longer than is necessary - so they already have a duty to delete it as soon as it is no longer needed. Considered that way, the right to be forgotten is simply putting the 'rights' side of an existing principle: allowing individuals to demand that those holding data fulfil their existing obligations.

This can form part of a bigger paradigm shift - a shift to a position where privacy is the norm rather than the exception, where the default is that individuals have choice (and to an extent power) rather than businesses or government bodies. This has many implications: in terms of browsing the internet, it should mean that browsing without being tracked would be the default, and tracking the exception. Tentative movements to address this have

already begun on both sides of the Atlantic, with the 'Do Not Track' initiative in the US [34] and the controversial and mislabelled 'Cookie Directive' in Europe, [35] and although neither had by the time of writing brought about much in terms of concrete results, they do both show a recognition that there is an issue to be addressed.

3.1 The right to delete

In the context of the holding of data, this paradigm shift could and should mean that the default concerning data should be that individuals do have the right to delete data connected to them, and that those that wish to retain data need to justify their holding rather than the reverse. The immediate corollary of this shift of assumptions would be the establishment of a general right to delete. That is, in general a data subject should be seen as having the right to delete any data held relating to them, and that those holding that data must put into place systems that allow that right to be enforced at any time.

How would this kind of right differ from the right to be forgotten which, as discussed in Section 1 above has many problems and would be very hard to establish? The first difference arises through the difference in names: calling it the right to delete rather than the right to be forgotten indicates a difference both in focus and in effect. The intention of the right should not be to allow people to erase or edit their 'history', but to control the data that is held about them. The change in name should make that purpose clear. Moreover, a 'right to delete' is a direct right - a right to act - whereas a 'right to be forgotten' appears to be a right to control someone else. This idea of control is connected closely with the association made between the idea of a right to be forgotten and restrictions on free speech, and on censorship. The change in name should help to make it clearer that the connection between a right to delete and censorship is tenuous at best - and in a practical sense non-existent.

3.2 Exceptions to the right to delete

The second and more important difference is in the use of the exceptions to the right, which set out when data should not be able to be deleted. The right to delete would be a qualified right - and those qualifications address the difficulties that appear inherent in the right to be forgotten. There are five principle categories of reason for which data might need to be preserved regardless of an individual's wishes to delete it - where the presumption should be in favour of retention rather than deletion.

1. Paternalistic reasons - where it is in the individual's interest that the data is kept, and society can override the individual's desire. The primary example of this is medical data;
2. Communitarian reasons - where it is in the community's interest that the data is kept. This might include criminal records, for example;
3. Administrative or economic reasons - where the economic or administrative needs of society require records to be kept. This could include tax records, electoral rolls and so forth;
4. Archival reasons - for keeping a good, accurate and useful historical record of events. This might include newspaper archives, blogs and so forth. This category is very important, but could easily be governed through a system by which a particular

database is agreed to be 'archival' in nature, and thus not covered by the right to delete - but also restricted in the uses to which it can be put and so forth. This is in itself another contentious issue. The British Library, for example, has campaigned for a 'right to archive', effectively asking for the right to archive web pages without needing to get permission from the website owners. [36] At first sight this might appear to be precisely the opposite of the shift of assumptions being suggested in this article, but in reality the two rights are quite compatible: both require close scrutiny and regulation of an archive. The British Library could be included on a 'register of archivists' that are permitted to keep such an archive - but required to control and report on that archive.

5. Security reasons - where the data is deemed to be needed for security purposes. This might include records of criminal investigations, or such communications records as are set out in data retention laws. This category is by its nature highly contentious, and should be subject to close scrutiny - including political scrutiny - and regularly reassessed.

These exemptions can be compared with the data protection principle of 'fair and lawful processing' concepts (consent, vital interests, administration of justice, functions of crown and public interest), 'processing exemptions' (research, history and statistics, and the special purposes exemptions: journalism, artistic use and literary use), and the exemptions to access rights set out in Schedule Seven of the Data Protection Act 1998. [37] All of these cover similar kinds of ground - so the concept of such limitations should be familiar and acceptable. Indeed, setting these terms out from a rights perspective could be part of a harmonisation process, making all these areas consistent and coherent. These exceptions should also deal with the key objections to the right to be forgotten as set out in Section 1.

The archival exceptions would prevent the right being used in any real way to 'rewrite' or 'erase' history - and allay the real fears of journalists that the right could be used to gag or censor them. Data is not synonymous with history: the right to delete could not be used to remove a record of where someone went to school, but it could be used to delete the record of what breakfast cereal they bought from an online supermarket or which websites they browsed one particular morning. The availability of the archival exception would depend not just on the nature of the data concerned but also the nature of the service or database in which it is contained. In terms of the school that someone went to, for example, the records held by the school itself or by the relevant local authority would be able to avail themselves of the archival exception, but a social networking site or similar kind of system would not. The function of Facebook's databases is not the maintenance of an accurate, useful historic record, but a current and potentially profitable social networking service. Another implication of this exception should be that what is already 'in the public domain' will remain in the public domain - although precisely what 'the public domain' consists of is something that will need to be regularly reassessed.

The other exceptions would deal with other objections to the idea of a right to be forgotten: the communitarian and paternalistic exceptions, for example, would remove the worries set out by Kenneth Clarke about the right causing problems for the portability of medical data or for legitimate information being used for credit histories. They would not, however, prevent a user from deleting records from Facebook that might be used inappropriately against them by potential employers, insurance companies or individuals with a grudge.

It should be specifically stated that 'supporting your business model' should not be a sufficient reason to deny data deletion - this could be viewed simply as taking data minimisation seriously, but needs to be explicit. One of the key purposes of rights is to spell things out so that people understand the principles, and might even begin to understand the reasons behind those principles. It is also important to ensure and remember that this kind of right should not be considered to be an alternative to the current ideas of data minimisation - nor does it remove any aspect of responsibility for data minimisation from the data processor or data controller - but should act as an additional safeguard, another level of protection for the individual.

3.3 Profiling and other derived data

The exemptions set out above cover the kinds of data for which deletion should not be possible - but there is another end of the spectrum: a category of data that would need to be specially highlighted as 'available for deletion'. That is, not just that the data subject has the right to delete them, but that attention must be drawn to them and it must be made simple, direct and clear how to delete them. The most direct example of this would be 'profiling' or 'channelling' data - so that an individual is able to delete information derived about them from their behaviour in one form or another. The reasons for highlighting this kind of data are twofold: firstly, because this kind of data can represent the most direct threat to people's autonomy, and secondly because profiling or derived data could be a way that data gatherers attempt to avoid or circumvent data minimisation rules in relation to the time that data is held. To give a simplified example, if someone searched for and looked at a particular website in January 2011, then if the periods of data retention suggested by the Article 29 Working Party in their Opinion on Search Engines are followed, [38] the fact that they performed that search could only be retained for six months, until June 2011. If at that point, however, whilst deleting that search log data the search engine provider creates some new 'profiling' data, categorising the person as a 'visitor of websites of that kind in early 2011', that profiling data could be classified as 'new' data in June 2011, and then kept for a further six months, before being incorporated into some new form of profiling data, and kept for another six months. Intelligent use of profiles can effectively extend data retention for unlimited periods - and hence special provision needs to be made to cover it.

3.4 Sensitive and non-sensitive data

When considering the right to delete, it is also important to consider the issue of the sensitivity of data. The idea that sensitive personal data should require more stringent conditions - and indeed a great many restrictions - is one that appears obvious, and the need for those holding it to provide justification for that holding is clear. The developing techniques of data aggregation and profiling mean that non-sensitive data also needs to be considered much more carefully. According to the rules set out in the Data Protection Act concerning 'sensitive personal data' [39] data concerning whether a person suffers from diabetes would be classified as 'sensitive personal data'. Data about whether the subject is a regular purchaser of sugar-free chocolate, or has ordered books about treatment for diabetes would not. Neither of these facts specifically indicates that the individuals concerned are diabetics - but if profiling is applied, even automatically, the

chances of the individuals being classified within categories that consist almost entirely of diabetics would be high. What is more, this example shows only the more obvious and intuitive kinds of connections that could be made, and that almost kind of 'sensitive' data can be inferred from what appears to be non-sensitive data. With detailed processing and large-scale data aggregation, even the most seemingly innocuous data, from sports followed or the kinds of news items read to choice of snacks or time of surfing on the internet can become highly significant. The data itself may not be sensitive personal data but is capable of revealing sensitive personal data. This has two direct implications: that the minimisation of even what appears to be non-sensitive data needs to be taken seriously, and that the right to delete should apply just as much to non-sensitive data as to sensitive data.

3.5 Deletion and anonymisation

There is another key issue in relation to the deletion of data - the issue of anonymisation. There is a close relationship between the two, and as and where it is technically possible the right of data deletion could be augmented with a form of subsidiary right - the right to have data anonymised. The primary right would be to delete data - but in certain circumstances a data controller could offer the option to anonymise the data instead, if the data subject would be willing to let that happen. The relationship between deletion and anonymisation is not a simple one. For one thing, it should be noted that if the right to delete is brought in, a data controller could avoid the possibility of that data being deleted by prior anonymisation - as the data would no longer be linked to an individual, no individual would have the right to delete it. Moreover, data is not always related to just one person - one clear example of this would be a group photograph in which a number of the people pictured are 'tagged'. That could bring a conflict of rights - if one person wants the data deleted but the other does not, whose rights have priority? Anonymity could apply here as well - in the photo example, it would be the tag that could be deleted rather than the photograph itself, effectively using the subsidiary right of anonymisation.

Even more importantly it must be remembered that anonymisation is far from a reliable process. Indeed, there is evidence to suggest that much supposedly 'anonymised' data can be 'de-anonymised', by combining it with other, often public, data sources. In 1997 Latanya Sweeney demonstrated that by combining an anonymised hospital discharge database with public voting records could produce identifiable health data. [40] Computer scientists have continued to work on de-anonymisation - their models are getting substantially stronger and more applicable to the kind of data now being generated on the internet. In a 2008 paper, Narayanan and Shmatikov of the University of Texas demonstrated by combining the databases of Netflix and the online movie database IMDB that if you knew the county someone lived in and one movie that they had rented in the last three years, they could be uniquely identified 84% of the time. Moreover, they suggested that their results could be generalised - and applied to most other similar databases. [41] Work in this field has continued - and its implications are significant. At worst, it can be argued that anonymisation is simply an illusion [42] - and even at best it means that it needs to be considered very carefully and its weaknesses taken seriously. [43]

3.6 The virtue of forgetting

As noted above, the idea of a right to delete is both nominally and qualitatively different from the right to be forgotten. Nonetheless, there are still aspects of forgetting that are closely related and both important and beneficial. Viktor Mayer-Schönberger has written compellingly about the virtues of forgetting in 'Delete'. [44] Perhaps most importantly in the context of this article, Mayer-Schönberger analyses how the developing 'default' of perfect digital memory takes control out of the hands of the individual, as their information and history becomes an indelible part of an mass of information usable and controllable by others. Moreover, it removes some of the positive effects of the passing of time. Digital memory can bring back information that has been forgotten for a reason, as part of the brain's method of navigating through life. As Mayer-Schönberger puts it:

> '... forgetting is not an annoying flaw but a life-saving advantage. As we forget, we regain the freedom to generalize, conceptualize, and most importantly to act.' [45]

Mayer-Schönberger's analysis is deep and detailed, providing strong arguments in favour of forgetting, and against the ideas presented by Bell and others that digital memory is a purely beneficial development. [46] Furthermore, Mayer-Schönberger has suggested a solution to the problem of 'excessive remembering' by digital systems, the idea of expiration dates on information - as he puts it, 'reviving forgetting'. His suggestion is an ingenious and interesting way to find solutions to the problem and in practice could have many benefits, though it also faces significant obstacles from both a technical and a business perspective. Establishing a right to delete could take it a step further, as it would put more control in the hands of the individual. The two could and should work together - implementation of expiry dates on certain forms of data would provide a kind of overarching control over data, while the specific right to delete would provide further autonomy and put further pressure on businesses to develop better, faster acting and more flexible business models.

4. Conclusions and implications of a right to delete

This article has set out some of the reasons that a right to delete is something that should be considered very seriously given the current state of affairs of the internet. It has suggested an approach to it that should be able to address the primary objections to the idea of a right to be forgotten. Rather than being an instrument of censorship, a restriction of freedom of expression or an attempt to erase or edit history, the right to delete can be seen as a change in the focus of data protection. The right to delete is a way to make data protection more about the rights and principles of data subjects and less about a legal framework for businesses to work around, as it currently often appears to be in practice. It would work as an extension and better implementation of data protection principles, first of all by extending data access rights. This could provide a boost for the concept of 'privacy by design': if the holder of data has to provide a means for a user to delete data, they must first provide fast and understandable access to that data, and to do this properly would mean taking data privacy into account right from the start.

There are significant barriers to be overcome before anything like a right to delete could become reality - in terms of 'workability' Kenneth Clarke may well be closer to the mark than he was in considering the potential problems of a right to be forgotten. However, if a way can be found for it to be implemented, the right to delete could have a very positive impact. It could give individuals the possibility of more control over their data and hence more autonomy. It could directly reduce the amount of data that is held - and hence that is vulnerable - as individuals exercise their right to delete. More importantly, it could force those holding data to justify why they're holding it - in such a way that the data subjects understand, for if data subjects cannot understand why the data is wanted, they might simply delete it. If there is benefit, and that benefit is made clear, why would an individual wish to delete that data? Most importantly of all, the fact that data could be deleted at any time could encourage the development of business models that do not rely on the holding of so much personal data.

This last point is perhaps the key to the next stage of development of the internet insofar as privacy is concerned. The amount of data removed by the direct exercise of a right to delete is likely to be insignificant compared to the reduction in data held as a result of any potential changes in business models, particularly if the right to delete is accompanied by equivalent shifts in terms of the gathering and processing of data. Over the last decade it has been the shift towards the business models of those such as Google and Facebook that has changed the face of the internet. If the next such shift is one that favours privacy and autonomy, that could be to the benefit of all.

---

[1] Paul Bernal is a Lecturer in Information Technology, Intellectual Property and Media Law at the School of Law, University of East Anglia, England.

[2] EC Communication to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, COM(2010) 609, p8

[3] See for example http://blogs.wsj.com/brussels/2010/11/30/the-right-to-be-forgotten/

[4] PINO, G. 2000. The Right to Personal Identity in Italian Private Law: Constitutional Interpretation and Judge-Made Rights. In: HOECKE, M. V. & OST, F. (eds.) The Harmonisation of European Private Law. Brussels: Hart. p237

[5] In France, beginning with its implicit recognition in article 40 of the 'Loi nº 78-17 du janvier 1978 relative à l'informatique, aux fichiers et aux libertés' and later in its effective inclusion as article 226-20 of the French Penal Code in 2000

[6] The online version of her article is at http://www.guardian.co.uk/commentisfree/libertycentral/2011/mar/18/forgotten-online-european-union-law-internet

[7] Commenting online on Mayes' article at http://www.guardian.co.uk/commentisfree/libertycentral/2011/mar/18/forgotten-online-european-union-law-internet

[8] Details of the event can be found at http://www.westminsterforumprojects.co.uk/forums/event.php?eid=235

[9] At University College Cork, CCJHR IV Annual Postgraduate Conference April 2010: Borders of Justice: Locating the Law in Times of Transition (http://www.ucc.ie/en/lawsite/eventsandnews/Conferences/ConferenceArchive/postGradConf2010/), BILETA 2011 (http://www.law.mmu.ac.uk/bileta/) and the Media@UEA symposium June 2011 (http://www.uea.ac.uk/ssf/media/symposium2011)

[10] His speech can be found online at http://www.justice.gov.uk/downloads/about/moj/our-ministers-board/speeches/clarke-speech-data-protection-260511.doc

[11] As set out in MURRAY, A. & SCOTT, C. 2002. Controlling the New Media: Hybrid Responses to New Forms of Power. Modern Law Review, 65, 491-516. and MURRAY, A. D. 2006. The Regulation of Cyberspace: Control in the Online Environment, Milton Park, Abingdon, UK ; New York, NY, Routledge-Cavendish.

[12] See for example BERNAL, P. 2011. Rise and Phall: Lessons from the Phorm Saga. In: GUTWIRTH, S., POULLET, Y., DE HERT, P. & LEENES, R. (eds.) Computers, Privacy and Data Protection: an Element of Choice. Brussels: Springer.

[13] Facebook eventually abandoned Beacon on 21st September 2009, following a settling a class-action lawsuit that had been brought in California accusing not only Facebook but a number of its allied retailers of breaching various US wiretapping and privacy laws. For an examination of the class action suit, see http://www.wired.com/threatlevel/2008/08/facebook-beacon/, and for the suit itself see http://www.wired.com/images_blogs/threatlevel/files/facebook_beacon_complaint0812081.pdf

[14] See for example the Google blog announcing the reduction at http://googleblog.blogspot.com/2008/09/another-step-to-protect-user-privacy.html

[15] Sony has acknowledged that 77 million users of Playstations and 25 million users who access the Playstation Network through PCs or Facebook may have had their data stolen. See http://www.soe.com/securityupdate/pressrelease.vm and http://www.soe.com/securityupdate/index.vm

[16] http://www.soe.com/securityupdate/pressrelease.vm

[17] See for example http://www.theregister.co.uk/2011/05/25/google_profiles_database_dump/

[18] See the report into the breach: POYNTER, K. 2008. Review of information security at HM Revenue and Customs. HM TREASURY. London: HMSO.

[19] See for example the ICO report into the affair at http://www.ico.gov.uk/~/media/documents/pressreleases/2010/ICO_statement_Street_View_01112010.ashx

[20] See for example this report into the UK government's potential acquisition: http://www.timesonline.co.uk/tol/news/politics/article7061114.ece and this into the German use of such data http://www.reuters.com/article/2010/02/07/germany-switzerland-bank-idUSLDE6160KO20100207

[21] LulzSec describe themselves as 'a small team of lulzy individuals who feel the drabness of the cyber community is a burden on what matters: fun'. As hackers, they have

claimed responsibility for a number of attacks, including one on Sony Pictures, and of bringing the CIA website offline for a period. See http://lulzsecurity.com/

[22] Downloadable from http://www.ico.gov.uk/~/media/documents/library/Data_Protection/Notices/response_to_moj_dpframework.ashx

[23] Details of the new penalties, and guidelines from the ICO as to how they are intended to be applied are available at: http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/ico_guidance_monetary_penalties.pdf

[24] See ICO press release on the ACS:Law penalty at http://www.ico.gov.uk/~/media/documents/pressreleases/2011/monetary_penalty_acslaw_news_release_20110510.ashx

[25] In an interview for simple-talk.com, at http://www.simple-talk.com/opinion/geek-of-the-week/ross-anderson-geek-of-the-week/

[26] In an interview for simple-talk.com, at http://www.simple-talk.com/opinion/geek-of-the-week/ross-anderson-geek-of-the-week/

[27] In an interview for simple-talk.com, at http://www.simple-talk.com/opinion/geek-of-the-week/ross-anderson-geek-of-the-week/

[28] ICO 2008. 'Taking stock, taking action'. London: Information Commissioner's Office. p6

[29] Reported for example in http://www.guardian.co.uk/technology/2010/jan/11/facebook-privacy

[30] Quoted for example in Wired, at http://www.wired.com/politics/law/news/1999/01/17538

[31] See http://technologizer.com/2010/05/11/facebook-privacy-fodder/

[32] Available online at http://www.legislation.gov.uk/ukpga/1998/29/schedule/1

[33] Working Party 'Opinion on data protection issues related to search engines', 00737/EN WP 148, Downloadable from http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2008/wp148_en.pdf

[34] See for example The Register's summary of the state of the 'Do Not Track' initiative in May 2011, at http://www.theregister.co.uk/2011/05/06/senate_do_not_track/

[35] The directive, labelled PE-CONS 3674/09, modifies Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (the 'ePrivacy Directive') and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws. It is available at http://register.consilium.europa.eu/pdf/en/09/st03/st03674.en09.pdf

[36] See for example http://news.bbc.co.uk/1/hi/technology/8535384.stm

[37] Data Protection Act 1998, s28, 29, 33, Sch. 7(9), Sch. 7(1), Sch. 7 (8) respectively

[38] Working Party 'Opinion on data protection issues related to search engines', 00737/EN WP 148, Downloadable from http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2008/wp148_en.pdf

[39] Data Protection Act 1998, Part I, Section 2

[40] SWEENEY, L. 1997. Weaving technology and policy together to maintain confidentiality. Journal of Law, Medicine and Ethics, 25, 98-110.

[41] NARAYANAN, A. & SHMATIKOV, V. 2008. Robust De-anonymization of Large Sparse Datasets. IEEE Symposium on Security and Privacy. 2008 ed.Available online at http://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf

[42] As suggested, for example, by Michael Colao at a meeting of the Society for Computers and Law in March 2011. See http://www.scl.org/site.aspx?i=ne19845

[43] See for example the work of Paul Ohm, in OHM, P. 2010. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. UCLA Law Review, 57, 1701-1778. Ohm analyses the work of computer scientists from Sweeney onwards and suggests that a full understanding of the weaknesses of the anonymisation process is required if methods to protect privacy are to be effective.

[44] MAYER-SCHÖNBERGER, V. 2009. Delete : the virtue of forgetting in the digital age, Princeton, N.J., Princeton University Press.

[45] MAYER-SCHÖNBERGER, V. 2009. Delete: the virtue of forgetting in the digital age, Princeton, N.J., Princeton University Press., p118

[46] In for example BELL, C. G. & GEMMELL, J. 2009. Total recall : how the E-memory revolution will change everything, New York, N.Y., Dutton.