



## Faculty Research Working Papers Series

### **Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing**

**Viktor Mayer-Schönberger**

John F. Kennedy School of Government - Harvard University

**April 2007**

**RWP07-022**

The views expressed in the [KSG Faculty Research Working Paper Series](#) are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.

**Useful Void**  
**The Art of Forgetting in the Age of Ubiquitous Computing**

Abstract

Viktor Mayer-Schönberger

As humans we have the capacity to remember – and to forget. For millennia remembering was hard, and forgetting easy. By default, we would forget. Digital technology has inverted this. Today, with affordable storage, effortless retrieval and global access remembering has become the default, for us individually and for society as a whole. We store our digital photos irrespective of whether they are good or not - because even choosing which to throw away is too time-consuming, and keep different versions of the documents we work on, just in case we ever need to go back to an earlier one. Google saves every search query, and millions of video surveillance cameras retain our movements.

In this article I analyze this shift and link it to technological innovation and information economics. Then I suggest why we may want to worry about the shift, and call for what I term data ecology. In contrast to others I do not call for comprehensive new laws or constitutional adjudication. Instead I propose a simple rule that reinstates the default of forgetting our societies have experienced for millennia, and I show how a combination of law and technology can achieve this shift.

# Useful Void

## The Art of Forgetting in the Age of Ubiquitous Computing

Viktor Mayer-Schönberger\*

In March 2007, Google confirmed that since its inception it had stored every search query every user ever made and every search result she ever clicked on. Google remembers forever.<sup>1</sup>

“хранить вечно“ (to be preserved forever) the KGB stamped the dossiers on its political prisoners. The Communist state would never forget the identity, believes, actions and words of those that had opposed it.<sup>2</sup>

Like the Soviet state, Google does not forget. But unlike the Soviet Union that ceased to exist fifteen years ago, Google has become an indispensable tool for hundreds of millions of people around the world, who use it every day.<sup>3</sup> We seem to have accepted that our digital society may forgive, but no longer forgets.

In this article I suggest that we should revive our society’s capacity to forget. The first part describes how we have unlearned to forget, due to a combination of technological innovations and the ensuing changing economics of information technology. In the second part I critique three conventional responses to our digital inability to forget. The third part reviews a non-traditional response of combining law and software put forward by Lawrence

---

\* Associate Professor of Public Policy, The John F. Kennedy School of Government, Harvard University. Version 1.01.

<sup>1</sup> Google to Anonymize Data, Wired Blog Network, March 14, 2007, [http://blog.wired.com/27bstroke6/2007/03/google\\_to\\_anony.html](http://blog.wired.com/27bstroke6/2007/03/google_to_anony.html) (last visited April 23, 2007).

<sup>2</sup> See LEV KOPELEV, TO BE PRESERVED FOREVER (1977).

<sup>3</sup> Verne Kopytoff, *Google surpasses Microsoft as world's most-visited site*, SAN FRANCISCO CHRONICLE, April 25, 2007, available online at <http://www.sfgate.com/cgi-bin/article.cgi?file=/c/a/2007/04/25/MNGELPF0DR1.DTL> (last visited April 26, 2007) (stating that more people use Google’s website than Microsoft’s, and that Google has the world’s most valuable brand), “It’s official: Google rules the world.”

Lessig. In contrast, I suggest a more modest proposal of reversing the data retention default, describe how such a proposal would work, and why it is superior to its alternatives. (For readers interested in my proposal only, please jump directly to part 4.)

## 1. The Demise of Forgetting

As humans we have the ability to remember. Yet, despite the remarkable capacity of our brain most of our memories fade over time. We forget. Fortunately, we tend to remember important things, and to forget less significant ones. Not only our individual memory fades. Because we are mortal each generation's collective memory would vanish if we fail to pass it on in time.

For thousands of years, humans have tried to preserve their memories beyond the life of each generation. Our ancestors have conveyed what they know to their children, thus keeping alive societal memory. They began to paint<sup>4</sup> and to write<sup>5</sup>. The printing press made retention cheaper and dissemination easier.<sup>6</sup> So did the spread of literacy.<sup>7</sup> Later, we added the phonograph<sup>8</sup>, the photograph<sup>9</sup> and film<sup>10</sup>. As the equipment to record and store became widely available people began to use it. More and more of what we said and did became preserved.

---

<sup>4</sup> See only the cave paintings at Altamira; PEDRO A. SAURA RAMOS, *THE CAVE OF ALTAMIRA* (1999).

<sup>5</sup> See the Sumerian cuneiform around the third millennium BC; Martha T. Roth, *Ancient Rights and Wrongs: Mesopotamian Legal Traditions and the Laws of Hammurabi*, 71 *CHI.-KENT. L. REV.* 13 (1995).

<sup>6</sup> ELIZABETH EISENSTEIN, *THE PRINTING REVOLUTION IN EARLY MODERN EUROPE* (2<sup>nd</sup> edition 2005); LUCIEN FEBVRE & HENRI-JEAN MARTIN, *THE COMING OF THE BOOK* (1976); FREDERICK G. KILGOUR, *THE EVOLUTION OF THE BOOK* (1998).

<sup>7</sup> MICHAEL E. HOBART & ZACHARY S. SCHIFFMAN, *INFORMATION AGES: LITERACY, NUMERACY, AND THE COMPUTER REVOLUTION* (2000); Eisenstein, *supra* note 6, at 92-98.

<sup>8</sup> MARK KATZ, *CAPTURING SOUND: HOW TECHNOLOGY HAS CHANGED MUSIC* (2004); WILLIAM HOWLAND KENNEY, *RECORDED MUSIC IN AMERICAN LIFE: THE PHONOGRAPH AND POPULAR MEMORY, 1890-1945* (2003).

<sup>9</sup> BEAUMONT NEWHALL, *HISTORY OF PHOTOGRAPHY: FROM 1839 TO THE PRESENT* (1982).

<sup>10</sup> DAVID PARKINSON, *THE HISTORY OF FILM* (1996); GEOFFREY NOWELL-SMITH, *THE OXFORD HISTORY OF WORLD CINEMA* (1999).

Yet, until recently conservation remained expensive, search was difficult and access limited. The costliness of continuous preservation forced us to carefully consider the trade-offs of retention and deletion. Only records viewed as important and valuable were kept. Libraries were expensive undertakings, requiring experts to keep the books in good order.<sup>11</sup> Our grandparents had their photo taken only at special occasions; and our parents would film us with Super 8 for short takes only, due to the cost of film.<sup>12</sup>

Searching our cultural artifacts of books and records, of photos and films, was a tedious and time-consuming task. Accessing and assembling a mosaic of a person's words and actions to create a biographical picture, for example, required training, dedication and expertise. Lay people rarely engaged in such ventures, leaving it to professionals to provide them with succinct summaries. Where government attempted to create huge data repository – as the Stasi tried in East Germany<sup>13</sup> – it often (but not always<sup>14</sup>) failed because of the inherent limitations to preserve, structure, and retrieve analog information. For centuries, our view of our past had been shaped by the technical and economic realities of information retention, access and retrieval.

Digital technology has changed this. The digital code has made processing as well as storing information not only easier but also cheaper.<sup>15</sup> Magnetic storage capacity has roughly doubled every year at constant cost. Similar increases in process power has made indexing our digitally stored content possible, thus offering almost instantaneous retrieval of vast amounts of stored information.<sup>16</sup> Finally, through digital networks – the Internet in

---

<sup>11</sup> MATTHEW BATTLES, *LIBRARY* (2004); see also the beautiful CANDIDA HÖFER & UMBERTO ECO, *LIBRARIES* (2006).

<sup>12</sup> J. HOBERMAN, *HOME MADE MOVIES: TWENTY YEARS OF AMERICAN 8MM & SUPER-8 FILMS* (1981).

<sup>13</sup> For a gripping account see the Academy Award winning movie “Das Leben der Anderen”.

<sup>14</sup> For a particularly tragic case of efficient use of government data see W. Seltzer & M. Anderson, *The Dark Side of Numbers: The Role of Population Data Systems in Human Rights Abuses*, 68 *SOCIAL RESEARCH* 481 (2001).

<sup>15</sup> NICOLAS NEGROPONTE, *BEING DIGITAL* (1995).

<sup>16</sup> Gordon E. Moore, *Moore's Law*, in RICHARD RHODES (ED.), *VISIONS OF TECHNOLOGY* (1999), at 243-244.

particular – we can access these information treasures around the world from our personal computers, cell phones and other digital devices.<sup>17</sup>

This has resulted in a drastic shift in our data retention behavior. For millennia it was difficult and costly to preserve. We would only do so in exceptional circumstances, and most frequently only for a limited period of time. For almost all of human history, most of what humans experienced was quickly forgotten. Today, however, retention of digital data is (relatively) easy and cheap. As a consequence, and absent other considerations, we keep rather than delete it. This is the central point: In our analog past, the default was to discard rather than preserve; today the default is to retain.

Credit bureaus store extensive information about hundreds of millions of U.S. citizens. Daniel Solove writes that the largest US provider of marketing information offers up to 1,000 data points for each of the 215 million individuals in its database.<sup>18</sup> We also see the combination of formerly disparate data sources. Solove mentions a company that provides a consolidated view at data from 20,000 different sources across the world.<sup>19</sup> It retains the data, he writes, even if individuals dispute its accuracy.<sup>20</sup>

Companies keep our air travel reservations on file even when we decide not to buy the ticket, together with rich information about us and our previous travel patterns.<sup>21</sup> Millions of cameras in public places – the UK alone is said to operate between 2 and 3 million<sup>22</sup> -

---

<sup>17</sup> MANUEL CASTELLS, *THE INTERNET GALAXY* (2001); MANUEL CASTELLS, *THE RISE OF THE NETWORK SOCIETY* (1996);

<sup>18</sup> See DANIEL J. SOLOVE, *THE DIGITAL PERSON – TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE* (2004), at 20; for an earlier account see SIMSON GARFINKEL, *DATABASE NATION – THE DEATH OF PRIVACY IN THE 21<sup>ST</sup> CENTURY* (2000).

<sup>19</sup> *Id.*, at 21.

<sup>20</sup> *Id.*

<sup>21</sup> See Edward Hasbrouck, *What's in a Passenger Name Record (PNR)?*, *The Practical Nomad*, <http://hasbrouck.org/articles/PNR.html> (last visited April 23, 2007).

<sup>22</sup> Timothy Zick, *Clouds, Cameras, and Computers: The First Amendment and Networked Public Places*, 59 *FLA. L. REV.* 1 (2007) at 15.

produce records of our movements that are kept.<sup>23</sup> Law enforcement agencies store biometric information about tens of millions of individuals even if these have never been charged with a crime.<sup>24</sup> Search engines retain each of our search queries, and keeps archival copies of our web pages long after we have taken them offline.<sup>25</sup>

This is only the beginning. With the advent of ubiquitous computing, of cheap GPS chips<sup>26</sup> in our cell phones, cameras and cars, of RFID tags<sup>27</sup> in everyday objects, and of tiny, networked sensors that surround us<sup>28</sup>, a more comprehensive trail of our actions will be collected than ever before. Given low cost of storage, ease of retrieval and potential value in accessing information, much of the data that is being collected will be kept for months if not years, as our societal default has shifted from deletion to retention.

This has drastic consequences beyond the obvious ability to know much more about other people's preferences, behaviors, actions and opinions than in the analog world of incremental forgetting. Living in a world in which our lives are being recorded and records are being retained, in which societal forgetting has been replaced by precise remembering, will profoundly influence how we view our world, and how we behave in it.

If whatever we do can be held against us years later, if all our impulsive comments are preserved, they can easily be combined into a composite picture of ourselves. Afraid how

---

<sup>23</sup> See Christopher Slobogin, *Camera Surveillance of Public Places and the Right to Anonymity*, 72 MISS. L.J. 213 (2002); Marc Jonathan Blitz, *Video Surveillance and the Constitution of Public Space: Fitting the Fourth Amendment to a World that Tracks Image and Identity*, 82 TEX. L. REV. 1349 (2004); Timothy Zick, *supra* note 22; see also JEFFREY ROSEN, *THE UNWANTED GAZE – THE DESTRUCTION OF PRIVACY IN AMERICA* (2000).

<sup>24</sup> See David Lazer & Viktor Mayer-Schönberger, *Statutory Frameworks for Regulating Information Flows: Drawing Lessons for the DNA Data Banks from other Government Data Systems*, 34 J. L. MED. & ETHICS 366, at 368, 371.

<sup>25</sup> This was uncovered when the US government subpoenaed search engines for search query data; see Katie Hafner & Matt Richtel, *Google Resists U.S. Subpoena of Search Data*, NEW YORK TIMES, January 20, 2006, <http://www.nytimes.com/2006/01/20/technology/20google.html?ex=1295413200&en=66c6a0f87da7e56d&ci=5090&partner=rssuserland&emc=rss> (last visited April 24, 2007).

<sup>26</sup> Evan Ackerman, *Geotagging Gets Tiny, Fast, Cheap, Oh Gizmo!*, March 21, 2007, <http://www.ohgizmo.com/2007/03/21/geotagging-gets-tiny-fast-cheap/> (last visited April 23, 2007).

<sup>27</sup> Radio-frequency Identification (RFID) permits automatic wireless identification of objects for tracking purposes.

<sup>28</sup> Mitchell Waldrop, *Pervasive Computing* (2003), available online at <http://wwics.si.edu/news/docs/pervcomp.pdf>

our words and actions may be perceived years later and taken out of context, the lack of forgetting may prompt us speak less freely and openly.<sup>29</sup>

This is the temporal version of a panoptic society, in which everything is being watched<sup>30</sup>; it is a society in which most of what is being recorded and collected is being preserved.

Regardless of other concerns we may have, it is hard to see how such an unforgetting world could offer us the open society that we are used to today.

## 2. Three Conventional Responses

Those that have access to and control over our personal data enjoy informational power. Concerns over such power (and its potential abuse) has prompted three types of reactions – the comprehensive legislative response, the constitutional reinterpretation response and the null response.

### a. Comprehensive Privacy Legislation:

Many privacy advocates argue that the comprehensive trail of personal digitized data that are retained requires a similarly comprehensive legislative reaction. While constraining data retention the goal of such legislative action is much broader. Only privacy statutes covering both the private and the public sector and encompassing all stages of the use of personal information - from collection and processing to retention and transferal - are seen as capable of containing and mitigating the danger to our privacy. So-called omnibus data protection is often bolstered with stringent auditing and enforcement procedures. The result is complex legal regimes that private and public sector users of personal information have to comply with.

---

<sup>29</sup> See also Jean Francoise Blanchette & Deborah Johnson, *Data Retention and the Panoptic Society: The Social Benefits of Forgetfulness*, manuscript, available online <http://www.rpi.edu/~blanc> (last visited April 20, 2007).

<sup>30</sup> Oscar Gandy has described the structures of our panoptic society, OSCAR GANDY, *THE PANOPTIC SORT: A POLITICAL ECONOMY OF PERSONAL INFORMATION* (1993); Michel Foucault has revived Jeremy Bentham's idea of the panoptic prison, MICHEL FOUCAULT, *DISCIPLINE & PUNISH: THE BIRTH OF THE PRISON* (1995); on Bentham's panoptic prison see also MIRAN BOZOVIC (ED.), *THE PANOPTICON WRITINGS* (1995).

In many industrial and post-industrial nations around the world, from Canada<sup>31</sup>, Argentina<sup>32</sup> and Chile<sup>33</sup> to Hong Kong<sup>34</sup> to Australia<sup>35</sup> and New Zealand<sup>36</sup> such legislation has been enacted, partially in response to public fears of large scale data collection and retention<sup>37</sup>; in Europe, the European Union (EU) Data Protection Directive, passed in 1995, obligates all twenty-seven member nations of the EU to pass stringent omnibus privacy laws.<sup>38</sup>

In nations where such comprehensive data protection regimes are still absent, like the United States, privacy advocates hope that media reports and general citizen unease over the threat to information privacy ultimately produce the ferment for political and legislative action.

At the same token, such a response is fraught with two substantial problems: political inertia due to collective action hurdles and potential structural overreach combined with limited actual impact.

---

<sup>31</sup> See Canada's Privacy Act (R.S., 1985, c. P-21 ) that applies to public sector use of personal information, and The Personal Information Protection and Electronic Documents Act (PIPEDA), 2000 c. 5 for use of personal data by the private sector.

<sup>32</sup> Argentina's Law for the Protection of Personal Data, Law No. 25.326.

<sup>33</sup> Chile's Law for the Protection of Private Life Ley Sobre Protección de la Vida Privada), Law No.19628 of August 30, 1999, Official Journal, August 28, 1999.

<sup>34</sup> Personal Data (Privacy) Ordinance

<sup>35</sup> Privacy Act 1988, Act No. 19 of 1988 as amended, available online at [http://www.privacy.gov.au/publications/Privacy88\\_100107.pdf](http://www.privacy.gov.au/publications/Privacy88_100107.pdf) (last visited April 24, 2007)

<sup>36</sup> Privacy Act 1993.

<sup>37</sup> For the political economy of privacy debates see COLIN BENNETT & CHARLES RAAB, *THE GOVERNANCE OF PRIVACY – POLICY INSTRUMENTS IN GLOBAL PERSPECTIVE* (2006); COLIN BENNETT, *REGULATING PRIVACY – DATA PROTECTION AND PUBLIC POLICY IN EUROPE AND THE UNITED STATES* (1992); DAVID FLAHERTY, *PROTECTING PRIVACY IN SURVEILLANCE SOCIETIES – THE FEDERAL REPUBLIC OF GERMANY, SWEDEN, FRANCE, CANADA, & THE UNITED STATES* (1989); Viktor Mayer-Schönberger, *Generational Development of Data Protection in Europe*, in PHILIP AGRE & MARC ROTENBERG (EDS.), *TECHNOLOGY AND PRIVACY: THE NEW LANDSCAPE* (1997), 219;

<sup>38</sup> 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, Official Journal November 23, 1995, L 281/31, available at [http://europa.eu.int/comm/internal\\_market/en/media/dataprot/law/index.htm](http://europa.eu.int/comm/internal_market/en/media/dataprot/law/index.htm)

Enacting privacy legislation has always been difficult to achieve politically. A relative small group of users of personal data, who stand to lose from any legislative initiative that constrains their activities, has every incentive to organize political opposition. Meanwhile, the millions of individuals who are potential beneficiaries of such privacy legislation are too diffuse and rarely energized enough to pressure legislatures to act.<sup>39</sup> In many ways, this is similar to public debates and resulting legislative action on intellectual property. There, too, a relative small, well-organized group of rights holders have been able to mobilize against legislative weakening and in favor of legislative strengthening of intellectual property rights, while the diffuse group of millions of consumers of intellectual property have failed to mobilize.<sup>40</sup>

Given the historical track record of privacy debates in the United States as well as the comparable trajectory of intellectual property debates (with a somewhat similar structure of actors on both sides) it is unlikely but not impossible for an event to stir up enough sustained public support for a comprehensive privacy act to get enacted.

The second problem with a conventional legislative response is structural. Laws are societal compromises frozen in time. They were enacted under particular circumstances and within a specific historical context. As new technologies emerge, economics change, and society evolves the delicate balance of values implicit in legal rules may become unhinged. To the extent such an imbalance energizes and mobilizes the citizenry, it may lead to an adjustment of the law. The fundamental aim of such corrective action is to re-establish the original balance.

Quite clearly, the change in retaining personal data brought about by digital technology may require legislative adjustment. Yet, what privacy advocates suggest – omnibus data protection legislation – is not a simple adjustment. Comprehensive privacy statutes create complex and far-reaching new rights, structures and processes, and fundamentally alter the way personal data can be used.

---

<sup>39</sup> MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION* (1971).

<sup>40</sup> See JESSICA LITMAN, *DIGITAL COPYRIGHT* (2001).

For example, under European data protection laws, generally personal data in general can only be collected, processed and stored with consent – and only for a particular purpose.<sup>41</sup> Any later repurposing of personal data requires the processor to obtain the individual's consent.<sup>42</sup> Individuals can require others to tell them what personal information is stored about them, for what purpose and from what source, and to which third parties (if any) the personal information has been passed on.<sup>43</sup>

While creating a complex new rights regime, in reality these rights are rarely exercised. Few request access to their data, and even fewer enforce their rights through legal action. In fact, over a ten-year period in Germany, not a single court case was brought by an individual claiming her privacy rights were infringed, even though Germany's data protection law shifts the burden of proof in asserting harm from potential plaintiffs to processors of personal data.<sup>44</sup>

In sum, omnibus data protection acts create completely new structures, processes, institutions, and rights that are rarely used. They introduce a fundamentally new regime to manage privacy claims, thereby replacing an existing societal value compromise with a new one. This is not to say that such a step towards a comprehensive information privacy regime ought not to be taken, but that it must be the result of a societal renegotiation of the underlying values. Given the legislative inertia I have described, it is unclear that our society is ready and willing to engage in such renegotiation any time soon.

#### b. Constitutional Reinterpretation:

We have long understood that some values are too important to be negated by simple majority rule. We protect them through constitutional norms, making it difficult to alter

---

<sup>41</sup> Art 7 EU Privacy Directive.

<sup>42</sup> Art 6 1.b. EU Privacy Directive.

<sup>43</sup> Art 12 a. EU Privacy Directive.

<sup>44</sup> VIKTOR MAYER-SCHÖNBERGER, *INFORMATION UND RECHT* (2001), 126; see also § 8 Bundesdatenschutzgesetz ([German] Federal Data Protection Law).

them. Instead our courts are called to interpret the meaning of Constitutional clauses protecting fundamental values in light of societal change. This requires care, as constitutionally protected rights are themselves finely balanced societal compromises over fundamental values frozen in time.<sup>45</sup>

One may suggest that the shift from forgetting towards remembering (and the ensuing retention of personal data that I have described) signifies precisely a situation in which the Supreme Court ought to reinterpret certain constitutional rights, especially if legislatures, perhaps due to the collective action hurdles I have mentioned, fails to act and thus leave constitutional protection undermined by our societal reality.

Arguments suggesting constitutional adjustment on First Amendment grounds are based on the notion that humans must have the right to engage in robust public debate. Limitless retention of one's utterances may preclude individuals from saying what is on their minds and impoverish public debate. It is the idea that the panopticon<sup>46</sup> causes a chilling effect that stifles free speech.

In *Laird v. Tatum*<sup>47</sup>, the Supreme Court was asked to adjudicate whether the Army's files containing antiwar activities of plaintiffs violates their freedom of speech.<sup>48</sup> Failing to identify concrete harm, the Court dismissed the case. Recently, the Court has warmed a bit to the chilling effect argument, but indicated that only action that is solely intended to chill speech could possibly violate the First Amendment.<sup>49</sup> This, however, will be difficult to show in most cases in which personal data of individuals is being recorded and retained.

---

<sup>45</sup> How the Supreme Court goes about its interpretive work, of course, is one of the central debates of constitutional theory. See SOTIRIOS A. BARBER ET AL, *AMERICAN CONSTITUTIONAL INTERPRETATION* (2003).

<sup>46</sup> See text and references *supra* note 30.

<sup>47</sup> 4098 US 1 (1972).

<sup>48</sup> For lower court decisions see *Donohoe v. Duling*, 465 F.2d 196 (1972); *Philadelphia Yearly Meeting of the Religious Society of Friends v. Tate*, 519 F.2d 1335 (1975).

<sup>49</sup> See *Meese v. Keene*, 481 US 465 (1987).

Christopher Slobogin has suggested another potential First Amendment argument, namely the ability to express oneself anonymously.<sup>50</sup> Over the last decades, the Supreme Court has struck down laws that required individuals to identify themselves or their affiliations before participating in public discourse.<sup>51</sup> Julie Cohen has argued that anonymity may be protected beyond protecting participation in public discourse.<sup>52</sup> While the Supreme Court has not yet decided a case of personal data collection and retention through the anonymity lens, one can conceive of cases that may prompt the Court to do so in the future. For example, recording individuals through public surveillance cameras – especially when they are engaged in communicative action - and automatically identifying them through face recognition is the kind of involuntary stripping of anonymity that the Supreme Court has been concerned about in the past.<sup>53</sup>

Fourth Amendment arguments focus on the act of collecting personal data. Once data collection can be characterized as Fourth Amendment “search”, such collection would require adhering to stringent procedural safeguards. In *Katz v. United States*<sup>54</sup>, the Supreme Court stated that in principle the bugging of a public phone booth was a type of “search”, which indicates that the argument is not without merit. But in the same case the Court held also that if one “knowingly exposes to the public” information, one cannot claim protection under the Fourth Amendment.<sup>55</sup> Subsequently the Court emphasized that it viewed mere observation without intrusion as not violating the Fourth Amendment.<sup>56</sup> Consequently, only when personal data is collected without the individual assuming that her words and actions

---

<sup>50</sup> Slobogin, *supra* note 23, at 257-258.

<sup>51</sup> *NAACP v. Alabama*, 357 US 449 (1958); *Talley v. California*, 362 US 60 (1960), *McIntyre v. Ohio Elections Commission*, 514 US 334 (1995); *Buckley v. American Constitutional Law Foundation*, 525 US 182 (1999).

<sup>52</sup> Julie Cohen, *A Right to Read Anonymously: A Closer Look at ‘Copyright Management’ in Cyberspace*, 28 CONN. L. REV. 981 (1996)

<sup>53</sup> Such automatic identification based on biometric parameters, like distances of facial features, is making significant progress; see Lamont Wood, *Feds: Accuracy of Face Recognition Software Skyrockets*, LIVE SCIENCE, April 13, 2007, available online at [http://www.livescience.com/technology/070413\\_face\\_recognition.html](http://www.livescience.com/technology/070413_face_recognition.html) (last visited April 24, 2007).

<sup>54</sup> 389 US 347 (1967).

<sup>55</sup> *Id.*

<sup>56</sup> *California v. Ciraolo*, 476 US 207 (1986); *Dom Chemical v. United States*, 476 US 227 (1986), *Kyllo v. United States*, 533 US 27 (2001)

are public, it is a potential “search” triggering Fourth Amendment constraints. Some have argued that the Supreme Court may need to more broadly re-conceptualize its interpretation of Fourth Amendment to respond to emerging privacy threats, but so far the Court has been reluctant do so.<sup>57</sup>

A third constitutional argument could be made based directly on the right to privacy. While this may seem the most obvious avenue, constitutional experts are reluctant.<sup>58</sup> The right to privacy is not explicitly mentioned in the Constitution; rather it has been “found” to be implicit in the First, Third, Fourth, Fifth, Ninth and Fourteenth Amendments.<sup>59</sup> Even after four decades of case history its contours remain vague, and the Court has been very reluctant to expand its reach beyond its (narrow) original confines.<sup>60</sup>

There is another important constraint limiting the practical usefulness of the constitutional reinterpretation approach: Constitutional guarantees, with very few exceptions, constrain only government, not private action. At best therefore constitutional reinterpretation may constrain governmental collection and retention of personal data, but will have no impact on the practices of the private sector. This may perhaps assuage libertarians solely worried about the power of the state, but leave unaddressed fears of a panoptic Google.

### c. The Null Response:

The third possible response is legal inaction, based on both procedural and normative considerations. Procedurally, if citizens are sufficiently concerned about comprehensive data retention, they will pressure legislatures to pass privacy legislation. The fact that (in the US at least) this has not happened indicates that our society is not troubled enough to want legislative action be taken. While those in the minority may lament the outcome, advocates of the null response may see it as democracy at work.

---

<sup>57</sup> See only Marc Jonathan Blitz, *supra* note 23.

<sup>58</sup> Slobogin, *supra* note 23, 263-267.

<sup>59</sup> *Griswold v. Connecticut*, 381 US 479 (using the term “penumbra”)

<sup>60</sup> PAUL M. SCHWARTZ & JOEL R. REIDENBERG, *DATA PRIVACY LAW* (1996) 36-39.

Moreover, the appropriate mechanism to address compelling minority concerns of constitutional adjustment is Constitutional adjudication. Even constitutional experts concede there is no constitutional right to societal forgetting. It is unlikely therefore that the current Supreme Court will find the general shift from forgetting to remembering violating a particular Constitutional guarantee.

If there is neither a willingness of the majority to pass legislation, nor a compelling normative case to be made for the institution adjudicating constitutional claims to intervene, inaction may be the most appropriate and efficient response compared to both a costly new regulatory regime and legal uncertainty caused by Constitutional re-interpretation.

The null response seems compelling, but the argument contains a potentially crippling weakness. The null response that citizens' preferences are adequately reflected in legislative outcomes is based on an idealized view of how modern democracies work. Political science theorists have long exposed and explored the importance of well-organized interests that are able to capture the legislative agenda, especially when these interests oppose a relative diffuse, only mildly activated majority. "Capture", collective action problem, issues of agency and interest politics diminish the ability of the majority to transform its will into law. To assume that these hurdles are absent in the context of data retention is naïve, and the societal shift from forgetting to remembering may not have been legitimated by the democratic process.

In addition, the null response is not necessarily the least costly. For example, given the evolutionary trajectory of technology, a small adjustment in the legal framework made today may be more economical than adopting a much more costly omnibus privacy act in the future.

It seems we are faced with three suboptimal responses: legislative action that introduces a comprehensive privacy framework much broader than the specific issue of retention of

personal data; limited constitutional interpretation that if at all may only cover government activities; or inaction despite an illegitimate societal shift away from forgetting.

### 3. Combining Law and Code - Lessig's Response:

Lawrence Lessig has offered a reason why conventional legislative measures are often failing in our digital world.<sup>61</sup> He suggests that human behavior can be constrained through different mechanisms. Law, which he dubs "East Coast Code"<sup>62</sup> is one of them, cultural norms and the market are two others.<sup>63</sup> In our digital age, a fourth mechanism has risen in importance: rules built into our information and communication technologies through software.<sup>64</sup> This Lessig calls "West Coast Code"<sup>65</sup>. As our digital technologies gain importance in our daily lives, he submits, software turns into a much more useful mechanism to influence and constrain our behavior than law.<sup>66</sup> Lessig thus proposes that effective regulation of our digital space cannot be achieved through law only, but requires the thoughtful combination of multiple mechanisms of constraint.

Lessig even applies his theory to information privacy.<sup>67</sup> He is cognizant of the legislative response, the omnibus privacy statute. But he also recognizes the significant weaknesses inherent in such a complex framework, including the lack of efficiency. In its place Lessig suggests a mechanism mix: enact legislation that grants individuals some kind of right over their personal information<sup>68</sup>, provide a technical infrastructure to trade personal information

---

<sup>61</sup> Lawrence Lessig, Code 2.0 (2006).

<sup>62</sup> Id, at 72.

<sup>63</sup> Id, at 122.

<sup>64</sup> Id, at 120-125.

<sup>65</sup> Id, at 72.

<sup>66</sup> Id, at 7.

<sup>67</sup> Lessig, *supra* note 61, at 223-232.

<sup>68</sup> In the first edition of Code, Lessig had suggested a property-like right; in his recent Code 2.0, heeding to criticism, he has switched to a system based on contract and liability. See LAWRENCE LESSIG, CODE (1999) 159-161; Lessig, *supra* note 61, 229-230.

at low transactional costs, and thus facilitate markets in which individuals can decide for themselves whether, under what conditions, and for what price they permit others to use their personal data. In essence, Lessig suggests that we create law and software – East Coast and West Coast Code – to provide effective and efficient protection for personal information.

Lessig hits an important point that advocates of omnibus privacy acts often disregard: Constraints on our behavior work best when they do not unduly burden legitimate behavior. This prompts him to suggest a system in which law protects claims over personal data, but lets markets negotiate their use – much like Coase has suggested for radio spectrum.<sup>69</sup> Technology enables these market transactions by lowering the cost involved in transacting so that the market functions efficiently. Technology also helps in enforcing, by making illegal behavior relatively more difficult than lawful one. Compared with omnibus data protection laws, Lessig’s proposal has the advantage of relative efficiency.

Lessig’s idea of combining law and software to regulate is not as radical as it may sound at first. In fact, Congress has regulated similarly in at least two domains of the information and communication realm: television content rating and copyright.

In the context of television content rating, Congress mandated that every television receiver built 2000 and after has to contain a special chip (the so-called V-Chip), which would be able to decipher special information sent with the ordinary television signal.<sup>70</sup> This information contains a rating of the content currently broadcast on that channel, based on seven dimensions (violence, language, sexual situations, dialogue, fantasy violence, drama, and crude humor) and six levels. Viewers (particularly parents) can set their TV so that only programs of a certain rating are shown. This system combines conventional laws with a technological infrastructure (the V-Chip etc), thereby ensuring that parents can utilize the (voluntary) TV parental guidelines effectively and at low cost.

---

<sup>69</sup> Ronald Coase, *The Federal Communications Commission*, 2 JOURNAL OF LAW AND ECONOMICS 1 (1959).

<sup>70</sup> 47 U.S.C. 303(x); see Jack Balkin, *Media Filters, the V-Chip, and the Foundations of Broadcast Regulation*, 45 DUKE L.J. 1131 (1996).

Copyright legislation offers a second case. Conventional copyright law prohibits the unlicensed use of copyrighted material except in certain relatively narrow circumstances. Due to advances in digital technology, copying and distribution of unlicensed works became easier and cheaper than obtaining a proper license. To remedy this situation, Congress could have facilitated a market infrastructure that would enable licensing at much lower cost – similar to Lessig’s idea of technology to support efficient markets for personal information. Instead, Congress in the Digital Millennium Copyright Act (DMCA)<sup>71</sup> decided to make unlicensed copying and distribution harder by prohibiting the tampering with technical mechanisms that protect copyrighted works. This constrains what kind of software can be developed and offered on the market. Those that want to violate copyright law will have more difficulty in obtaining the technical tools to do so. Illegal conduct is becoming more expensive relative to legal one and thereby changes how users behave.

In sum, the feature of Lessig’s privacy proposal to combine law and software is useful in theory and applicable in practice.

Yet, in the context of this article – to conceive of a useful response to the societal shift from forgetting to remembering – his proposal suffers from a weakness not dissimilar to omnibus data protection legislation. It is too broad. It suggests we create a full-fledged market for personal information with the necessary technical, legal, and societal infrastructure. It assumes that people will engage and behave rationally in these markets. It is a complex and far-reaching proposal to respond to what is a much narrower issue: the need for our society to remember to forget. Much like omnibus data protection acts Lessig’s proposal may be too substantial for policymakers and citizens alike. (To his defense, Lessig crafted his proposal as a comprehensive response to the privacy threat in cyberspace, not as a reaction to the data retention issue.)

---

<sup>71</sup> Pub. L. No. 105-304, 112 Stat. 2877 (codified as amended at 17 U.S.C. § 512 (1998)). See *United States v. Elcom Ltd.*, 203 F. Supp. 2d 1111, 1127-29 (N.D. Cal. 2002) (upholding DMCA to constitutional challenge).

What can we learn from Lessig's proposal for our case? We need a focused, effective response to our societal inability to forget. It does not require the setup of a comprehensive privacy regime, while combining law and software may help make our response effective.

#### **4. Reverting to the Default - Reintroducing Forgetting into our Society:**

Where we used to forget over time, we now have the capacity to perfectly remember. Technological change has enabled us rather than erase data to retain data by default, often for a very long time. This is particularly troubling when this data pertains to us, to our lives, actions, beliefs, and preferences.

I propose that we shift the default when storing personal information back to where it has been for millennia, from remembering forever to forgetting over time. I suggest that we achieve this reversal with a combination of law and software. The primary role of law in my proposal is to mandate that those who create software that collects and stores data build into their code not only the ability to forget with time, but make such forgetting the default. The technical principle is similarly simple: Data is associated with meta-data that defines how long the underlying personal information ought to be stored. Once data has reached its expiry date, it will be deleted automatically by software, by Lessig's West Coast Code.

This may sound either simplistic or radical (or both), but I believe it is neither, as I hope you agree when you come to understand how I envision it to work, and when I explain its advantages and shortcomings.

##### **a. The Workings of Forgetting:**

Technically the proposal is quite modest. We already manage and store an ever-increasing amount of meta-data about our data. An abundance of meta-data drives much of what we call Web 2.0. The file systems of our modern operating systems manage meta-data on all of our files. Applications and appliances add meta-data without much intervention: Digital

cameras store a wealth of meta-data with each picture taken, our photo management software adds another set of it. Our music library software manages meta-data for our music, from links to lyrics and album art and playlists, to our own tags and preferences. And yes, Digital Rights Management (DRM) provides (and enforces) meta-data of rights for many of the copyrighted digital works we have stored on our devices.<sup>72</sup> Soon, we'll see geographical meta-data be added to our media and used to visualize our travels and enable us to connect to others in our geographic vicinity. My proposal simply adds another type of meta-data: information about data life expectancy.

In most instances of individual use, little would change from the world we know. We would configure our word processing and spreadsheet software once, so that our documents are automatically tagged to have an almost infinite life (although there would be easy ways to change the expiry date).

The situation would begin to look differently for software managing webcams and surveillance cameras. There the legal mandate would require software code to set a relatively short default for storage, perhaps a couple of days or weeks.

My proposal would also alter how Internet Cookies<sup>73</sup> would work. Cookies already have an expiry date, but that date is often set decades into the future, thus in effect creating data that our Internet browsers never forget. When coding cookies software engineers would have to ask in earnest how long the cookie data should actually persist, and set it accordingly.

Google and other search engines may have to change their practices as well. No longer would they be able to store search queries forever. They would have to be deleted – forgotten – over time. Amazon would have to adjust by changing its software to automatically “forget” what books you bought years back. (As a result customers might

---

<sup>72</sup> See Viktor Mayer-Schönberger, *Beyond Copyright: Managing Information Rights with DRM*, 84 DENV. U. L. REV. 181 (2006).

<sup>73</sup> Their correct name is “Persistent Client State HTTP Cookies”; see RFC2109 – HTTP State Management Mechanism, <http://www.ietf.org/rfc/rfc2109.txt>; see also Viktor Mayer-Schönberger, *The Internet and Privacy Legislation: Cookies for a Treat?*, 1 W. VA. J. L. & TECH. 1 (1997).

perhaps experience surging accuracy in Amazon’s recommendations. After all, whose literary preferences stay constant over years?)

The change would also affect cell phones. Currently cell phone software often stores traffic data – for example the last ten calls made – forever. I propose that stored traffic data in cell phones, too, is tagged with an expiry date of say a week or two, after which the cell phone software automatically deletes that piece of traffic data from your cell phone’s memory. The same would apply to photos and videos on cell phones, as well as personal information in cell phone address books, albeit with a much longer time frame of perhaps months or years.

Don’t get me wrong: I do not suggest that your cell phone erases its memory every other day. What I suggest is that by default digital personal information is not kept forever. Think of it as the analog to our non-digital world, where we jot down temporary notes on little pieces of paper, which are easy to throw away (or lose). In the world of post-its and napkins, the default of forgetting (or at least losing access to) is quite obviously built in.

My proposal would also affect the trajectory and use of future technologies, especially in the areas of networked sensing. Much has been written about the potential privacy threat of pervasive digital sensors.<sup>74</sup> By limiting the duration such sensor data can be retained, we could address a significant part of the challenge.

For some companies, the technical changes mandated by my proposal may require work, even adjustment of their business practices. For others it may already reflect their privacy commitments. Take Microsoft, for example. Its own exemplary “Privacy Guidelines for Developing Software Products and Services” state that “[a]ny User Data stored by a company should have a retention policy that states how long the data should be kept and the manner in which it should be removed from all data stores.”<sup>75</sup>

---

<sup>74</sup> See *supra* note 28 for references.

<sup>75</sup> Privacy Guidelines for Developing Software Products and Services, Version 2.1, p. 9, available online at <http://www.microsoft.com/downloads/details.aspx?FamilyId=C48CF80F-6E87-48F5-83EC-A18D1AD2FC1F&displaylang=en> (last visited April 24, 2007).

My proposal aims to reintroduce the concept of forgetting over time into our digital realm. My goal is to shift the default back from retaining forever to deleting after a certain time. At its core, my proposal does not envision that users cannot change the expiry dates if they want to – the digital equivalent of taking a napkin of somebody’s address and putting it neatly in a file to preserve it. In a sense, we have had the ability to do that almost forever. But it required a deliberate act; and that is what I suggest our digital devices require from us, too.

b. A Spectrum of Options:

Legislatures could – if they wanted to go beyond the minimum of what I suggest – mandate that organizations that store personal information only use software that can manage expiry meta-data (likely on a record rather than file level) and that they keep the appropriate expiry data up-to-date.

This might force credit bureaus and direct marketing database vendors to think a bit harder about their data retention policies. It might force the operators of computerized reservation systems in the travel industry to delete records after the reservation was cancelled, and not retain them for much longer as is currently their practice. It may force government agencies to ponder whether and when to delete citizen data, rather than keeping it stored, just in case.

Policymakers seriously concerned about citizens’ privacy in the digital age could even go one step further. They could compel organizations and companies that process particularly sensitive personal data to digitally sign the expiry meta-data and - taking a page from the DMCA’s anti-circumvention legislation<sup>76</sup> - prohibit the unauthorized alteration or tampering of such meta-data. This could make an unauthorized change in the expiry data – for example by a third party – more difficult. I am not proposing such a legislative measure, but it would be entirely consistent with my proposal.

---

<sup>76</sup> Codified as 17 U.S.C. § 1201 (2000) et seq.

It could even be applied to photos and videos. Envision small “permission devices” that we carry with us like key rings. They could be set to “permit” or “deny”, and would transmit that response wirelessly when queried. A new generation of digital cameras would poll the permission devices of the person recorded. The camera would then set the expiry meta-data accordingly, say a couple of days or weeks in case somebody’s device was set to “deny”, an much longer if all permission devices were set to “permit”. The camera could even provide a visual cue whether somebody’s device said “deny”.

My proposal offers a range of options to legislators, a spectrum from the core idea to more complex, and far-reaching versions. Where policymakers find their specific legislative sweet spot is up to them. Moreover, no option promises perfection. Leakages – meta-date being changed to circumvent deletion - will persist, albeit of varying degree. But perfection is not my aim, changing the default is.

c. Advantages and Weaknesses:

My proposal is modest on at least three dimensions:

- Technologically, as I have described, it utilizes much of the ideas, infrastructures, and mechanisms that already exist. In most cases, it requires only limited technical modifications.
- Legally, it introduces to the digital realm the default against retention that is so familiar to us in our analog world, and thus reflects the fundamental value compromise that underlies our Constitution. It does not create new rights, or new institutions. Instead, it simply resets the default to where we as a society had negotiated it to be. The proposal is also modest in combining law and software. As I have explained, this, too, is nothing novel or radical.

- Politically, it is less controversial than omnibus privacy acts. Like other combinations of law and privacy enhancing technology (PET)<sup>77</sup> its general idea is well understood. Focused on a well-defined and specific issue it likely is not overbroad. It does not overly burden entire business sectors, yet will change how we perceive privacy in our society. And it equips us well for a world of ubiquitous computing and pervasive sensors.

This proposal also has the advantage of eminent practicability over other alternatives. We know that omnibus data protection acts have not prompted individuals to exercise the new privacy rights they have been given. Constitutional reinterpretation is if not unlikely uneven, while the null response may make an eventual legislative reaction extremely costly. Finally, unlike Lessig's comprehensive privacy infrastructure, my proposal does not require markets in personal privacy to emerge and does not rely on the rationality of human beings as market participants.

Some may dislike the proposal because it does not solve all challenges of information privacy. This is true. But it is not intended to. The proposal is focused on a very specific issue: the default of data retention. Changing that default will have significant positive effects on privacy, but it is no privacy catchall.

Others may criticize my use of Lessig's theory of East and West Coast code, and point out the inherent difficulty in shaping society by shaping technology. This point is well taken. I do not suggest to foresee how users may respond to the switch in retention and deletion defaults, how they perhaps may undercut the legislative intention, or pervert it. I am fully cognizant that technology is not an exogenous force that is impressed on society, but rather evolves from a delicate dance with society, in which each side is influencing the other.<sup>78</sup> Therefore, attempts to influence societal behavior through the shaping of technology may fail.

---

<sup>77</sup> Herbert Burkert, *Privacy Enhancing Technologies: Typology, Critique, Vision*, in PHIL AGRE & MARC ROTENBERG (EDS.), *TECHNOLOGY AND PRIVACY – THE NEW LANDSCAPE* (1997), 125-141.

<sup>78</sup> See e.g. SHEILA JASANOFF ET AL (EDS), *HANDBOOK OF SCIENCE AND TECHNOLOGY STUDIES* (1995).

What I am suggesting is a legislative mandate that technology conform to our real-world practices, rather than shape society in a completely novel way. The necessary technological tools are mostly developed, and do not require innovation at a level that may spur an utterly unforeseen dynamic with society. In sum, the chances of success of my modest proposal, while surely not certain, are significantly higher than other proposals of law shaping technology to influence society.

In one important aspect, though, my proposal is radical: It is radical in its desire to change when, what and how we remember and forget in the digital realm. In this aspect and over the long term it may have significant consequences for how our digital spaces are shaped.

## **5. Conclusions:**

For millennia, humans have had to deliberately choose what to remember. The default was to forget. In the digital age, this default of forgetting has changed into a default of remembering. As these digital memories make possible a comprehensive reconstruction of our words and deeds even if they are long past, it may constrain our willingness to engage in and further our open society. Three conventional reactions to this perceived threat have been put forward: the legislative, the constitutional and the null responses. Each of them has significant structural weaknesses, limiting its value. Lawrence Lessig has suggested combining law and software to create a novel response to the novel threats in the digital age, including threats to our informational privacy. His premise is a good one, but his privacy proposal may be too broad.

In this article I have suggested an alternative proposal: to reintroduce the concept of forgetting over time into our digital realm. My aim is to shift the default back from retaining forever to deleting after a certain time. I have described the legal and technical components of my proposal and how they would work together. I have suggested an implementation spectrum based on how far policymakers and the public want to go to ensure forgetting.

Evaluating my proposal relative to the alternatives I have argued that it is relatively modest, on a number of implementation dimensions, making it comparatively easy to be adopted, yet it is radical enough to facilitate the fundamental shift from remembering to forgetting that is so central to our society's fundamental values.