Data retention and the panoptic society: The social benefits of forgetfulness*

Jean-François Blanchette
Science and Technology Studies
Rensselaer Polytechnic Institute

Deborah G. Johnson
School of Public Policy
Georgia Institute of Technology

Contact information:
Jean-François Blanchette
Centre de recherche en droit public
Faculté de droit, Université de Montréal
C.P. 6128, succ. centre-ville
Montréal (Qué), Canada, H3C 3J7
Tel: (514) 343-6111 x1760
Fax: (514) 343-7508
Web: www.rpi.edu/~blanc
Mail: blanc@rpi.edu

Running title: Data retention and social forgetfulness

Abstract: Modern information systems not only capture a seemingly endless amount of transactional data, but also tend to retain it for indefinite periods of time. We argue that privacy policy must address not only collection and access to transactional information, but also its timely disposal. One unintended side-effect of data retention is the disappearance of social forgetfulness. Social forgetfulness allows individuals a second chance, the opportunity for a fresh start in life. We examine three domains in which social policy has explicitly recognized the importance of such a principle: bankruptcy law, juvenile crime records, and credit reports. In each case, we frame the issue not solely in terms of individual privacy protection, but rather in terms of the social benefits of forgetfulness. We examine how different policy approaches to privacy might handle the retention of data and propose a general framework for constructing retention policies. The broad conclusion of the paper is that data retention and disposal should be addressed as a fundamental characteristic of information, not in piecemeal fashion, or as an afterthought.

Keywords: Data retention, social memory, social forgetfulness, informational privacy, privacy policy, surveillance.

*This paper has greatly benefited from comments of participants at the "ACM Policy '98 Conference," organized by the Association for Computing Machinery, the "Computer Ethics: a Philosophical Enquiry (CEPE'98)" conference, held at the London School of Economics, and the "Graduate Student Conference on Technology and Identity," held at Cornell's Science and Technology Studies department, as well as from comments by Robert Gellman, David Charmichael, Daniel Poulin, Ann Cavoukian, and three anonymous reviewers. The first author has also benefited from the support of a doctoral grant from Canada's Social Sciences and Humanities Research Council.

> — It is not enough to keep repeating that memory is socially structured. To have come so far invites a further step. The next thing is to discover what qualities of institutional life have distinctive effects on remembering. Mary Douglas, *How institutions think.*
>
> — Cheerfulness, the good conscience, the joyful deed, confidence in the future — all of them depend, in the case of the individual as of a nation, on the existence of a line dividing the bright and discernible from the unilluminable and dark; on one's being just as able to forget at the right time as to remember at the right time; on the possession of a powerful instinct for sensing when it is necessary to feel historically and when unhistorically. This, precisely, is the proposition the reader is invited to meditate upon: the unhistorical and the historical are necessary in equal measure for the health of an individual, of a people and of a culture. Frederic Nietzsche, *On the uses and disadvantages of history for life.*

On December 28, 1997, Swiss cellular phone users were distraught to learn that the position of their phones (within a few hundred meters) was automatically and continuously registered by their service provider, Swisscom. While this is an inevitable feature of cellular telephony (in order to forward a call to a particular user, service providers must first ascertain the position of the phone with respect to the network), what made this revelation particularly disturbing from the privacy standpoint was the fact that Swisscom retained the data for a duration of *six months to a year and half* (AP 1997).

This incident is paradigmatic of a problem that has been largely overlooked in the privacy literature to date: control over personal information is not only effected through selective access, but also through selective *retention* of such information. That is, control is not only a question of who has and who does not have access to personal information (nowadays, seemingly everyone but its producer), but who gets to retain or discard it. Most privacy commentators focus on access and control, and address retention only as an afterthought—if at all. A central concern of this paper is to make the importance of this component explicit: we argue that data retention must figure as an important element of any comprehensive account of informational privacy.

We begin by framing the data retention issue within broad concerns over the lack of privacy protection in modern democratic societies. Secondly, we place the issue in the context of a tension between the importance of institutional/public memory and forgetfulness. Once the issue is framed as such, we go on to examine three domains of life in which the idea of the "fresh start"(where individuals move on, leave their past behind them and begin anew) plays a role. We then examine how different approaches to privacy policy—regulatory, market-based, and technology-based—deal with data retention. We conclude by sketching a framework that provides a more comprehensive approach to the issue.

1. PUTTING THE DATA RETENTION ISSUE IN CONTEXT

An enormous literature now documents concerns about and threats to personal privacy arising from new information and communication technologies. Concern heightens each time new technologies give rise to new forms of data collection. In the 1990s attention has been focused especially on transactional data (web browsing, credit card use, intelligent highways), by contrast with the 1970s and 1980s when concern was with the scale of record-keeping and collection of personal data. We will not describe the practices or technologies that give rise to such concerns, as an abundant literature already documents this, as well as the privacy policies extant in many countries. Most recently the European Union has become a focus of attention as it struggles with the harmonization of privacy policies of EU countries and with transborder data flows to non-EU countries.[1]

[1]. See Schwartz and Reidenberg (1996) for an extensive review and analysis of this question with regard to the United States.

We agree with others who have suggested that the apparatus of a *panoptic* society is slowly, but surely, being put into place in the U.S. (Gandy 1993) Democracies are generally thought of as societies in which individuals have a high degree of individual liberty and government power is limited and checked. Yet, it appears that information and communication technologies are moving us rapidly toward a panoptic society. The panopticon is Bentham's prison environment, as described by Foucault (1975), in which prison cells are arranged in a large circle with the side facing the inside of the circle open to view. The guard tower is placed in the middle of the circle so that the inside of each cell is in plain view of the guards. The amount of data currently collected as we go about our everyday lives—intelligent highway systems, consumer transactions, traffic patterns on the Internet, medical, educational, financial, and insurance records, and so on, strongly suggests we are moving into a panoptic society. Even if the data is not collected by a single, Orwellian-like entity, but rather by a mixture of public and private institutions, and even if what is observed is not necessarily amalgamated into a single dossier, the possibility of synthesis remains. Clearly, such a panoptic society presents fundamental challenges to the exercise of democratic freedoms and responsibilities.

Again, most of this is not new and we will not belabor the point. Rather we want to draw attention to the fact that most of the work that has been done on this issue has focused almost exclusively on how to control access to data (and the corresponding value of privacy), and neglected retention (and the corresponding value of social forgetfulness). Data protection policies have not proceeded from any comprehensive analysis of the problems occasioned by data retention. Instead, sector by sector, decisions have been made regarding the length of retention of data, with little attention being been paid to the cumulative effect of these piecemeal decisions.

Our approach to data retention begins from the insight that the endurance of data is a feature that has invisibly but powerfully changed with the shift from paper-and-ink to electronic systems of record-keeping. In the paper-and-ink world, the sheer cumbersomeness of archiving and later finding information often promoted a form of institutional forgetfulness—a situation with parallels to human memory.[2] The forgetfulness of the paper-and-ink world was implicit in the material being of institutions, the available storage space, the budget for file cabinets, etc. Often the institution's memory/forgetfulness was not even recognized as a policy issue but dealt with as a matter of physical facilities.[3] In many cases, as storage technologies have gained in practicality and lowered in price, the shift to an electronic medium changed the default position from one of forgetfulness to one of memory.[4]

Whether the paper and ink environment or the electronic environment favors data retention, the point remains that decisions about length of retention of data (institutional memory) may be made unintentionally or in an ad hoc manner, rather

---

[2]. This is somewhat echoed by the European Directive on Data Protection, which extends its protection only to cases where "the processing of ... data is automated or if the data ... are contained ... in a filing system structured according to specific criteria relating to individuals, so as to permit easy access to the personal data in question." (European Directive 1995)

[3]. Of course, retention policies are influenced by a variety of factors beyond the availability of archiving technologies, most notably fear of litigation and regulatory requirements.

[4]. In some respects, though, data may well endure longer in paper form than in an electronic environment, depending on a variety of factors. As David Charmichael, County Records Manager and Archivist for Westchester County in New York, testifies, "Westchester County still retains its first book of records from 1684, but its first computer tapes from 1977 are unreadable after just 21 years." (Charmichael 1998) In other words, institutional memory can turn out, in an electronic environment, to be a function of how often and what kind of technology changes an institution makes. When new technology is accommodating, data endures and it takes an intentional act to delete it, whereas when new technology is not accommodating, data may become effectively unusable.

than with an eye to privacy policy or institutional memory per se.[5] We find ourselves in a world that captures endless data on us and then decides (sometimes by failing to decide) how long to retain this data. When data endures, institutional forgetfulness shrinks and some important values may shrink with it—values which are fundamental to democratic society. In other words, we must ask, what are the social implications of a lack of institutional forgetfulness?

We will begin our investigation of this question within the U.S. context, for several reasons. First, there is a general consensus that, in the U.S., too little is being done to stop the onslaught of personal data collection. There is even, to some extent, a consensus on the nature of the problem in the U.S. It is that privacy protection policy has been ad hoc and piecemeal, rather than comprehensive (Regan 1995; Gellman 1997). At the same time (and perhaps ironically), the U.S. has traditionally understood itself to be a place where individuals could get a "second chance." The idea that an American citizen can sometimes "wipe the slate clean" and start anew is, no doubt, tied to the immigrant, pioneer histories of so many Americans.[6] Whatever its origins, the idea is in tension with current U.S. data collection and retention policies.

The idea that Americans value the opportunity for a "fresh start" was recognized in the early literature on privacy, and periodically recurs in current literature. Westin and Baker (1972), in their seminal work, *Databanks in a Free Society*, understood that this value was perceived to be under siege because of computers:

> "Many citizens assume, out of a variety of religious, humanistic, and psychiatric orientations, that it is socially beneficial to encourage individuals to reform their lives, a process that is impeded when individuals know (or feel) that they will automatically be barred by their past 'mistakes' at each of the later 'gate-keeping' points of social and economic life. Because the computer is assumed not to lose records, to forward them efficiently to new places and organizations, and to create an appetite in organizations for historically complete records, the computer is seen as threatening this forgiveness principle." (Westin and Baker 1972, p. 267)

Interestingly enough, Westin and Baker went on to point out that the key question about erasure or non-circulation of derogatory information was *not* a technical matter in the organizations they visited. It was an issue of social policy, on which society has to choose between the "forgive-and-forget" and "preserve but evaluate" theories of record-keeping in each substantive area (p. 268). In his study of police surveillance practices, Gary Marx has underlined how surveillance information "transcends time," that is, "it is available for analysis many years after the fact, and in totally different interpretive contexts." (Marx 1986, p. 150) He remarks how this threatens to undermine some basic American values:

> "The idea of 'starting over' or moving to a new frontier is a powerful concept in American culture. The beliefs that once a debt has been paid to society it

---

[5]. In fact, this is the conclusion the reader is forced to make when reading Schwartz and Reidenberg's (1996) survey of American data protection law: all requirements for retention of data are requirements of *minimum* duration, motivated by administrative requirements. In their analysis, Schartz and Reidenberg place great faith in the need for institutions to divest themselves, for reasons of efficiency, of the burden of accumulated data, thus enacting an ad hoc institutional forgetfulness, but also acknowledge that marketing divisions may well wish to keep the data, in order to establish long term consuming patterns. (Schwartz and Reidenberg 1996, §10-1(a)(4), §10-2(a)(4), §11-1(a)(4), §11-2(a)(4), §12-1(a)(4), §12-2(a)(4), §13-1-(a)(4), §13-2(a)(4))

[6]. This is echoed in Frederick Turner's classic thesis on the American frontier ideal, *The Idea of the Frontier in American History* : "In the long run, the effective force behind American democracy was the presence of the practically free land into which man might escape from oppression or inequalities which burdened them in the older settlements." (Turner [1920] 1986, p. 274)

> is forgotten and that people can change are important American traditions. Americans pride themselves on looking at what a person is today rather than what he may have been in the past. Devices, such as sealed or destroyed records, prohibitions on certain kinds of record keeping, and consent requirements for the release of information, reflect these concerns. However, with the mass of easily accessible files, one's past is always present, for erroneous or sabotaged data, as well as for debts that have been paid. This can create a class of permanently stigmatized persons." (Marx 1988, p. 223)

Of course, the extent to which Americans truly have valued, or continue to value, the opportunity to move on beyond one's past (especially when it is weighed against other goods, such as law enforcement) is an open question. By contrast with Westin and Baker, and Marx, Gandy (1993) has more recently articulated the value of forgetfulness, but with a more defensive thrust. Referring to "the right to be forgotten" as one of the fundamental principles of data protection identified by Flaherty (1989) in his study of privacy policies in Western industrialized societies, Gandy explains:

> "[t]he right to be forgotten, to become anonymous, and to make a fresh start by destroying almost all personal information, is as intriguing as it is extreme. It should be possible to call for and to develop relationships in which identification is not required and in which records are not generated. For a variety of reasons, people have left home, changed their identities, and begun their lives again. If the purpose is non-fraudulent, is not an attempt to escape legitimate debts and responsibilities, then the formation of new identities is perfectly consistent with the notions of autonomy I have discussed." (Gandy 1993, p. 285)

But, while Westin and Baker, Marx and Gandy, and yet others have drawn attention to the value of starting over, of having a portion of the past forgotten, the issue has been cast, implicitly or explicitly, as one involving *a tension between personal or individual privacy and social goods*. They have portrayed the issue as a matter of balancing individual privacy against such social goods as law enforcement, government efficiency, or national security. Yet, there is reason to believe that this framing of the problem is inaccurate and biased against individual privacy.

The lesson of the 1980s and early 1990s is that when personal privacy is put into a cost-benefit analysis, it generally loses. The needs of government agencies and private organizations or institutions—for more accurate and efficient information systems so as to further their goals (law enforcement, national security, administrative efficiency) overpower the desire (need, interest, or right) of individuals to have information about them kept private. Regan (1995) describes how this framing of the issue has led to the loss of privacy protection in several major public policy contexts. She argues against such a reductive framing of privacy on grounds that it does not recognize the *social* importance of personal privacy. Hence, in our analysis of institutional forgetfulness, we want to argue for forgetfulness as a social good, not just an individual good.

2. THE VALUE OF SOCIAL FORGETFULNESS

Privacy as an individual good and privacy as a social good are inextricably tied. To see this, one need only appreciate that the kind of world we live in makes us into certain kinds of beings and certain kinds of beings are essential for a certain kind of world. For example, democracy depends on individual citizens who are capable of formulating plans for their lives, taking action, thinking critically and making decisions. Yet, individuals of this kind can not develop in an environment of constant surveillance. The problem is not just that democracy is squelched when

individuals live in fear of repercussions for any non-conforming behavior: it is also that the mere fact that one is being watched changes the way one behaves, as Bentham and Foucault have taught us. Individuals change their behavior when they believe they are being watched, and come to see themselves as they believe they are seen by their watcher. The very nature of self and the kind of personalities that develop in a surveillance society are different.[7]

The argument is thus an argument for privacy both as an individual good and as a social good. Privacy is not just something individuals want because it makes them feel good or is good for them; rather, privacy is good for society insofar as it promotes the development of the kinds of individuals who are essential for democracy. A world in which there is no forgetfulness—a world in which everything one does is recorded and never forgotten—is not a world conducive to the development of democratic citizens. It is a world in which one must hesitate over every act because every act has permanence, may be recalled and come back to haunt one, so to speak. Of course, the opposite is equally true: a world in which individuals are not held accountable over time for the consequences of their actions will not produce the sense of reponsibility that is just as necessary to a democratic society. Thus, achieving the appropriate degree of social forgetfulness is a complex balancing act, ever in tension between the need to hold accountable, and the need to grant a "fresh start."

In order to begin understanding the requirements of retention policies, we have examined three policy arenas in which forgetfulness seems to play an important and explicit role: bankruptcy law, juvenile crime records, and credit reporting.[8] Bankruptcy law involves civil law, juvenile crime records, criminal law, while the regulation of credit reporting is more concerned with private institutions. We have examined these domains to find out if the apparent forgetfulness in these policies is real, to learn how forgetfulness was understood in the development or implementation of each policy, and to get a better overall feeling for how the tension between memory and forgetfulness has been played out in American social policy. We have also examined the arguments in these domains with an eye to re-deploying them in other domains and help us construct a comprehensive approach to data retention.

2.1 Bankruptcy Law

The first thing to note about bankruptcy law is that the discussion surrounding it does, indeed, recognize forgetfulness (and forgiveness) as a social good. In the first pages of a 1989 study of bankruptcy and consumer credit in America, the authors write:

> "Bankruptcy is a powerful phenomenon. It is financial death and financial rebirth. Bankruptcy laws literally make debts vanish. When a judge signs a paper titled 'Discharge,' debts legally disappear." (Sullivan, et. al. 1989, p. 4)

And later:

> "At the heart of all bankruptcy law, for individuals and for businesses, is the discharge of debts and other legal obligations, the 'fresh start.' The notion of beginning anew, of rebirth, lies near the center of our restless, westward-moving culture and is also the central proposition of its dominant

---

[7]. See Reiman (1995) for a recent and lucid articulation of this argument.

[8]. There are of course several other mechanisms within law concerned explicitly with mediating the tension between social justice and the opportunity to start over, e.g., free pardon, remission of sentencing, amnesty, statutes of limitations, etc. The precise makeup of such devices is naturally highly dependent on the social mores of the times: in France and Britain, for example, free pardon proved a useful mechanism to increase the size of both royal armies and new colonies. (Foviaux 1970)

> religions. Whether a bankrupt debtor, given more time, can pay in full or can pay little or nothing, the relaxation of strict legal obligations is the indispensable centerpiece of American bankruptcy law." (p. 20)

Of course, the textbooks on bankruptcy law and historical accounts of the development of these laws also make it clear that bankruptcy serves the interests of creditors as well as debtors:

> "... bankruptcy law is a supercollection device for creditors. Indeed, American bankruptcy law arose from two separate bodies of English law, one designed to protect debtors and the other to aid creditors. ... ordinary debt collection law has serious flaws from a creditor's point of view. Its two most important weaknesses are that it is purely state law, making collection across the country very difficult; and it is competitive, with every creditor for itself. Bankruptcy law immediately captures all the debtor's assets in one country-wide net after a single filing. It also restrains actions by any individual creditor, permitting creditors to act collectively, often through a trustee, to preserve asset values and to ensure a fair distribution." (p. 20)

While the literature we examined did express the concern for forgiveness for mistakes and the good of letting individuals move on, there are reasons to believe that these values alone would not have led to the forgiveness of bankruptcy, were it not for the fact that creditor interests were also served by the forgiveness. Moreover, government (social) interests were at work insofar as there was a perceived need to respond to periodic national financial crises and to facilitate individuals (especially those involved in business) in getting back into economic activity (Warren 1935).

The literature on the history of bankruptcy law supports Regan's thesis insofar as it describes a tension between individual and institutional interests which was finally (and perhaps, only) resolved when there was a coming together of *institutional interests* (creditors' interest in a non-competitive way to obtain whatever they could), *individual interests* in being able to start afresh (having their mistakes forgiven and forgotten), and *social interests* (in responding to major economic crises and getting entrepreneurs back into the economy).

Our research on bankruptcy law thus supports the idea that Americans recognize a social good of forgetfulness. Moreover, the research supports Regan's conclusion that arguments in favor of social forgetfulness (and privacy protection in general) are more likely to succeed when they are cast in terms of a social good rather than purely in terms of individual interests.

2.2 Juvenile Crime Records

Juvenile justice has evolved considerably over the last few centuries, concurrently with changing social conceptions of both children and the role of the State. Although there are many different and competing visions of how the State should intervene with regard to juvenile crime, one prominent train of thought has been the liberal (progressive) view of the State as protector of juveniles. Such a view primarily aims at rehabilitating juveniles through de-emphasizing their offences, and highlighting their treatment needs (Guarino-Ghezzi and Loughran 1996). Judge Mack powerfully echoes the sentiments underlying the liberal view:

> "Why is it not the duty of the state, instead of asking merely whether a boy or girl has committed a specific offense, to find out what he is, physically, mentally, morally, and then if it learns that it is treading the path that leads to criminality, to take him in charge, not so much to punish

> as to reform, not to degrade but to uplift, not to crush but to develop, not to make him a criminal but a worthy citizen." (Mack 1909, p. 107)

Of course, any such goal of rehabilitation must be carefully reconciled with other principles of justice, such as punishment and offender accountability. Juvenile justice statutes, both in the United States and in England, clearly indicate how the courts are expected to hold a balance between the protection of the public and that of the individual child. Section 1 of the Uniform Act states as one of its goal:

> "... consistent with the protection of the public interest, to remove from children committing delinquent acts the taint of criminality and the consequences of criminal behaviour and to substitute therefore a program of treatment, training and rehabilitation." (Parsley 1978, p. 182)

However, the public interest is here not only defined in terms of protection from delinquent elements, but also in terms of a "reserve capital", that is, the need to safeguard society's future. Not only has society an immediate interest in protecting itself from criminal elements, but in the case of juvenile delinquents, it has a future interest in preventing "the deprived and delinquent children of today from becoming the deprived, inadequate, unstable and criminal citizens of tomorrow" (Bean 1981, p. 126). Clearly, the State has much to gain in avoiding the huge social and economical costs that follow from committing individuals, from an early age, to a lifelong relationship with criminal justice.

Note that such a rehabilitative program is congruent with a number of different philosophical views on the nature of juvenile crime, (and the concomitant views with regard to the most appropriate form of punishment). Whether one holds that a child's criminal behavior is truly criminal or rather simply "naughty", whether she is held competent or not to understand the consequences of her actions, it is nevertheless understood that, following a certain purgatory, a young person's mistakes should not unduly burden her future goals:

> "... for those offences that could be called 'crimes' a child should not be expected to have a criminal record for behaviour that may be transient or reflect a particular stage of development." (Bean 1981, p. 131)

This is the justification for the special provisions within juvenile crime statutes aimed at removing the stigma of a juvenile court history. For example, the Code of Virginia includes provisions

> "... for the automatic expungement of juvenile records, for offences that would be felonies if committed as an adult, at the age of 29. All other offences may be expunged at age 19, if five years have elapsed since the juvenile's last contact with court. ... an individual may petition for expungement of all records pertaining to his/her case after 10 years since the date of the last hearing in juvenile court." (Virginia 1996)

There is thus recognition of the value of social forgetfulness embodied in policies on juvenile crime records. However, echoing our previous discussion on bankruptcy, it is important to note that the background discussion of these provisions point to a coming together of social and individual interests. Individuals are allowed to move on beyond their juvenile criminal records not just because it is good for them, but as well because society has an interest in turning juvenile offenders into law-abiding adults.

2.3. Credit Reports

Consumerism, as a way of life, would be all but impossible without the availability of consumer credit. Without it, families simply could not afford the

houses, cars, appliances, and electronic gadgets nowadays synonymous with the good life. The credit reporting industry has grown out of the desire for businesses to maximize opportunities for consumers to acquire such goods and services, while attempting to exclude those likely to default on their loans. As James Rule explains, "the art and science of credit management lie in determining, in advance, who will pay and who will not, and in screening credit applicants accordingly" (Rule 1973, p. 178).

Credit evaluation is based on the simple principle that past actions provide a good indication of future behavior. Credit bureaus thus seek to acquire the most complete information possible on individuals, so that their clients (businesses, credit-lending institutions, insurers) may make the most educated guess possible about whether or not to extend credit to applicants. Far from being limited to financial information, the reports assembled by credit bureaus may contain information relating to convictions, suits, employment history, past addresses, family status, etc. In fact, before regulators stepped in, almost any information that could be legally obtained was seen as fair fodder for the credit bureaus' files, but most importantly,

> "... credit bureaus place a special emphasis on seeking unfavourable or 'derogatory' information. ... it is much more efficient to aim at excluding bad risks than at including good ones, and derogatory information is to this extent at a premium." (Rule 1973, p. 193)

Thus, with regard to our previous discussions of bankruptcy and crime records, credit bureaus' activities would seem to go directly against the idea of granting the opportunity for a fresh start. Such past blemishes are *precisely* what the credit bureaus are paid to look for:

> "Worst of all, in the eyes of the credit grantors, are bankruptcy petitions, since they indicate a desire to shirk all debts, which is the most serious sin of all in an industry which profits only from willingness to pay." (Rule 1973, p. 194)

In the 1960's, more and more people availed themselves of the services of credit reporting agencies, and for an ever-widening range of purposes. The potential for abuse grew to the point that Congress felt compelled to regulate this booming industry through the Fair Credit Reporting Act (1971, revised 1997).[9] The Act was designed to cover a broad range of issues with regard to the activities of credit bureaus; its stated purpose was to protect individuals from the deleterious effects of credit reports, by establishing precise rules under which personal information can be reported. Most pertinent to our discussion, it defined certain categories of information that are subject to obsolescence: bankruptcies, suits and judgments, paid tax liens, accounts placed for collection or charged to profits or loss, and records relating to a crime. For each category, the Act established precise time limits after which information must be deleted from credit reports.[10] The FCRA thus ensured that the social forgetfulness principles established in the case of bankruptcy and juvenile crime records were not overwhelmed by the new data collection and aggregation practices of credit bureaus.

---

[9]. See McNamara (1973) for a more complete legislative history of the Act.

[10]. Even within those rules, credit bureaus found ample room to gnaw at the forgetfulness principle: "[I]n *Equifax, Inc.*, an FTC administrative law judge found that the reporting agency violated the Act by inserting phrases in its reports such as, "[i]n compliance with the Fair Credit Reporting Act, no additional information can be reported from this former employer concerning employment experience prior to seven years ago." The quoted phrase was inserted in consumer reports only when Equifax believed it had *adverse* obsolete information." (Sheldon 1994, p. 160)

In fact (perhaps inadvertently), the Act went even further. It prohibited the reporting of "any other adverse item of information" predating the report by more than seven years. It also omited to make clear not only what it meant by "item of information," but also how, and from what point in time, it should be judged "adverse." This is problematic since, as one analyst noted, "'Items' may well be continuing matters, such as divorce proceedings or, in investigative reports, disputes with neighbors or employers." (Willier 1971, p. 55) The interpretive flexibility afforded by such loose formulation, combined with fears of non-compliance with the Act, would seem to naturally force upon credit bureaus a conservative reading of what legislators sought to include within the category of "adverse information":

> "Since what may be adverse to one creditor, insurer or employer may not be adverse to another, absent any uniform and objective criteria for judgment, almost any items of information must be treated by the agency as adverse. In the extreme, this includes places and time of residence. ... In short, a consumer reporting agency should look upon *any* item of information as adverse for purposes of the seven years rule." (p. 55)

That is, except for the special categories mentionned above, the Act essentialy limited credit bureaus to a memory of seven years or less. Were it not for the generous conditions under which these obsolescence rules may be altogether skirted, the FCRA might have thus provided for some of the strongest policy in current legislation to implement a right to be forgotten.[11]

Despite its implementation flaws, the FCRA clearly represents a continuation of the philosophies outlined in the case of bankruptcies and juvenile crime records. If the judicial system has sought to provide individuals with some (if limited) means to unburden themselves from their past, the FCRA extends these policies to the new threats posed by data collection, aggregation, and reporting. [EXTEND: why is the FCRA about *social* forgetfulness!!!!]

3. THE NEW THREATS TO SOCIAL FORGETFULNESS

From the vantage point of business, credit reports highlight how personal information is most useful in aggregate form and accrues in value through accumulation over time. More broadly, the case of credit reports points to a changing social conception of personal information and privacy: concurrently with (and in spite of) the rise of privacy concerns, personal information has come to be seen as a valued commodity, an essential element of modern business needs (Davies 1997). Nowhere is this more evident than in the case of transaction-generated information (TGI), which records the details of our interactions with organizations or indviduals (phone calls, purchases, geographical location, banking transactions), facilitating aggregation and inordinately increasing our capacity for social memory. As is often the case with computerization, there is *in principle* nothing fundamentally new about TGI; rather, it is both the scope of and the new possibilities offered by the enterprise that promise to alter social memory in both subtle and dramatic ways:

— *Quantity:* as more and more of our activities are taking place over electronic networks, more categories of data are being collected every day. From an initially fairly limited set including phone calls, banking and credit card transactions, the

---

[11]. The rules limiting retention are waved under conditions easily met by almost any substantial credit, job, or insurance application. As a manual from the *Associated Credit Bureaus* explains: "Congress accepted the argument of some 'specialty' consumer reporting companies who make reports on consumers where large sums are involved, and exempted certain reports from the obsolescence section and any adverse item, no matter how old, may be reported if the report is being done for a credit transaction or life insurance policy which will be for at least $50,000; or for employment purposes where the annual salary will be at least $20,000." (Associated Credit Bureaus 1975, p. 710)

list now includes highway tolls, e-mail, Web browsing, cellular phones, grocery shopping, etc.[12]

— *Granularity:* for each category of transactions, a finer granularity of data collection is possible; a phone call over a cellular network may be recorded in terms of originator, destinator, duration, time of day, type of device used for the call, geographical location of device, movement of device during the call, network services used, etc. This increased capacity for precise metering of user's activities is part of the tremendous attractiveness of TGI for organizations.

— *Cross-correlation:* once collected, TGI is easily aggregated and correlated with other kinds of data: web browsing, demographics, credit card transactions, and cellular use together provide a much finer resolution of the digital persona than each can by itself.

— *Predictive power:* most importantly, quantity plus diversity plus cross-correlation combined lead to the possibility of "discovering" information not (explicitly) present in the data collection process itself. In other words, such data has *predictive power*. Because data is collected in electronic format, it is easily amenable to a variety of treatments: multidimensional and statistical analysis, neural networks, information discovery systems, all technologies precisely aimed at extracting new information from the vast warehouses of electronic information gathered by organizations. Even when the information is not available in a suitably discrete format, image analysis software or text analysis algorithms may be used to extract pertinent data from video or free-flowing texts. Such technologies may be used with regard to marketing, network management, credit-risk analysis, sales productivity, etc., with the hope that they may help discover rules, patterns of behavior, and predict the future with some reasonably good probability.[13]

While critics of the panoptic society have justly remarked on the *ubiquitousness* of data collection practices, we underline how such practices invisibly extend the *persistence* of social memory and diminish social forgetfulness. What the above list points to is a subtle yet dramatic change in the nature of this memory. Human activities and interactions which were not part of the public record have now the possibility of being registered in varying levels of details. In most cases however, there seems to be little concern over the effects of data retention. In fact, organizations have come to see and use such transaction-generated information as a legitimate and highly useful competitive asset.

With this in mind, we now examine how various privacy policies deal—explicitly or implicitly—with the idea of social forgetfulness and how they can be applied to a comprehensive approach to data retention.

5. A COMPREHENSIVE APPROACH TO RETENTION

How then are policy makers to approach the issue of determining retention periods for data? The issue is complex, as one must manage the ever present tension beetwen accountability and forgetfulness. Furthermore, little thought has been given so far to the question. Nevertheless, even a rough framework would be more useful than the current neglect that fails to recognize that a significant quality of democratic society is imperiled by the blind and automatic retention of information, transactional and otherwise.

---

[12]. Although not yet quite of the same nature, videotaping of public spaces will eventually also fall within this category, especially when coupled with face recognition technology (Thomas 1998). In the UK alone, an estimated 200,000 cameras cover public spaces. (Davis 1997, p. 150)

[13]. For a more detailed discussion of the technologies of data mining and knowledge discovery, see Mattison (1996).

We will not attempt to precisely and exhaustively classify data within various categories, assigning to each some specific retention period. While some researchers have developed such an approach with regard to the sensitivity of personal information (Wacks 1989, p. 226ff), such attempts have generally been abandoned in recognition that their dependence on evolving cultural norms makes them ineffective outside of their local contexts. The OECD guidelines thus remark that "...it is probably not possible to identify a set of data which are universally regarded as being sensitive." (OECD 1980) Similarly, the European Directive recognizes the difficulty of establishing any precise criteria of sensitivity, apart from a few categories which are held to be universally (at least within Europe) problematic:

> "Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life." (European Directive, art. 8.1)

It seems similarly unlikely that there exist any set of universally adequate principles from which to determine specific retention periods for specific categories of data. The best one can do at this stage is to establish some general categories of information that ought to fall under specific guidelines, and leave the adjudication of precise retention periods to local data protection authorities. In areas where well-established mechanisms exist to adjudicate the retention of personal information (e.g., credit reports and the obsolescence rules of the FCRA), we will "merely" reaffirm the need to consider social forgetfulness. In areas where no such mechanisms exist (such as TGI), we will attempt to establish some preliminary approaches to forgetfulness. Our model incorporates insights from both regulatory and technological approaches to data protection and retention.

5.1 The spectrum of retention

A rough caracterization of personal information and records may be obtained by using the idea of a *spectrum of retention*, from *permanent* documents to information not recorded at all. That is, at one end of the spectrum of retention are "permanent" documents, such as birth records; Next follows documents attached to individuals for (usually) lengthy periods of time, such as criminal, credit, or medical records; The next category contains information that maintains relevance for shorter periods of time, such as commercial transactions falling under the seven years rule, ; Lastly, at the other end of the spectrum is data that is never collected, has no permanence, is destroyed immediatly after being used. For each category, we can establish the following preliminary guidelines, as regard to retention policy:

— *Permanent documents:* certain documents never leave the individual, not even after his death.[14] Thus, birth or death certificates do not really fall under our discussion.[15]

— *Long-term records:* Next is data that is kept for (usually) significant periods of time, such as criminal, credit, or medical records. In this case, the actual framework seems to be well established. If the actual degree to which forgetfulness is valued varies with the times, institutional mechanisms with which to debate the issue exist: sealed records for juvenile crime records, judicial pardon for adult offenders, obsolescence rules for credit reports, all provide mechanisms whereby social

---

[14]. Although this is by no means self-evident: Ian Hacking (1990) provides an account of the historicity of such collections practices, which grew together with the expansion of the rational State in 17th-century France, England, and Germany.

[15]. Roger Clarke (1994) examines the history of the birth certificate as the "base or root document" of the identification documents "pyramid," and the various problems that ensues, mostly that "[w]ithout evidence to connect a person with the person named in the birth certificate, the certificate establishes nothing about that person."

forgetfulness is achieved, if not in practice, at least in principle. We will thus not concern ourselves with this category of information, beyond reaffirming the value of social forgetfulness for democratic society.

— *Medium-term records:* this forms the bulk of the records concerning individuals—electronic transactions, banking information (7 years or less), etc, etc. [I AM NOT SURE WHAT GOES HERE YET] Records managers use a three-tier categorization of records...

— *Flash records:* Data may never need be collected in the first place. At trip to the convenience store does not, at present, establish records, nor does it need to. In fact, as more and more privacy analysts have suggested, transactions that leave no records are desirable, possible, and could be the norm, without any threat to security:

> "[m]ainstream activities have led to a presumption on the part of many organizations that they need to have identity in order to conduct almost any kind of transaction. There remain many circumstances, however, in which the identity of the person with whom an organization performs transactions is of no consequence. In fact, the majority of the transactions undertaken between individuals, and between individuals and corporations, are still conducted anonymously." (Clarke 1994)[16]

What contentious, of course, is to decide which transactions may or may not be conducted in this way. What needs, and what does not need to be recorded?

We will not be concerned, obviously, with permanent documents. For long-term documents, we do not need to reinvent the whell here, only reaffirm the need to carefully adjudicate between accountability and social forgetfulness. Retention is also not an issue for the last case, but which type of transactions fall into this category is contentious. Our main concern is thus with medium-term information. We now examine which mechanisms (legal, regulatory, technical) are available to achieve social forgetfulness.

5.2 Achieving forgetfulness

What means are available to achieve forgetfulness, and which is best used for each category that we have identified.

— *Anonymisation:* In the case of data that is collected, but eventually anonymised, such as medical data. There are two problems with the idea of anonymising data: one is that it can be used to form judgments on groups (Vedder, Schreuders, and Van Kralingen 1998), which may eventually become detrimental for individuals identified with the group. Thus, the application of knowledge discovery tools to anonymised data may be detrimental if (a) the tool may eventually enable the re-identification of the individual (b) the tool may create identification of characteristics at the level of the group, which may be equally detrimental for the individual. Vedder, Schreuders, and Van Kraligen point out that privacy legislation is hopelessly ill-equipped to deal with such a type of what they call *categorial privacy violation* .

---

[16]. Phil Agre (1997) locates the source of that presumption at the level of a fundamental of tendency of computer science, the conflation of representation and reality: "The mirror metaphor, as well as the conceptual and linguistic tendency to conflate records with the entities they represent, makes it seem reasonable to employ a 'natural' identifier, such as name and birth date, as the primary key. But, ... a cold look at the technical requirements for a primary key provides a compelling case for the use of 'surrogate keys'—arbitrary numbers used solely to signify an entity instance's existence and not its other attributes." (Agre 1997, p.51)

The other problem is that while anonymising data may seem like an atractive solution to the problems posed by retention, it's implementation is not straightforward. As one researcher states:

> "Organizations often release and receive medical data with all explicit identifiers, such as name, address, telephone number and social security number, removed on the assumption that patient confidentiality is maintained because the resulting data look anonymous. However, in most of these cases, the remaining data can be used to reidentify individuals by linking or matching the data to other databases or by looking an unique characteristics found in the fields and records of the database itself." (Sweeney 1997, p. 98)

Thus, anonymising of data is not necessarily an alternative to erasure of data, as the European Directive seem to imply: "Data may be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected..." (European Directive 1995, art. 6.1.e)

— *Regulatory mechanisms:* The 1978 French "Informatique et libertés" law states, under article 28, that:

> "[u]nless otherwise provided for by the law, information may not be stored in personal form beyond the period stated in the application for an opinion or in the declaration, unless such storage is authorized by the Commission." (Flaherty 1989, p. 180)

Every data collection must be authorized by the French data protection authority. In its authorization, the CNIL specifies the maximum duration period for which the data may be kept, in accordance with the needs of the collection and regulatory requirements of other French legal bodies (archival requirements, for example). Again, the wording of the law leaves open the possibility for mere anonymisation of data.

— *The technological approach:* Data that may need to be collected, but does not need to be correlated to other data. This is the realm of Chaum's digital pseudonyms (Chaum and Evertse 1987; Chaum 1981; 1985; 1992)

The recent German law over telecommunication services effectively blends the regulatory and technological approach: technological devices must be *designed* so as to collect as little data as possible:

> "The design and selection of technical devices to be used for tele[communication] services shall be oriented to the goal of collecting, processing and using either no personal data at all or as little data as possible." (IKG 1997, art. 2, §3.4)

— *The bureaucratic approach:* Retention schedules, which specify the retention periods for institutions. Such schedules are built on a set of criterion that does not, usually, include "social forgetfulness", but rather, is built with an eye to avoid litigation, and conform to legal requirements for retention of records and financial statements. In a way, one can say that retention schedules have been influenced by technological possibilities, legal requirements, lately by marketing needs, but not generally by the need for social forgetfulness. Thus, there already exists extensive classifications of data with regard to their retention periods, in the form of retention schedules. For example, a number of statutes, both in the private and the public sector, govern retention of data. In the public sector, retention and archiving policies dictate which portion of the government's administrative, judicial, and legislative activities must be preserved and for how long: the *General Records*

Disposal Schedules of the Government of Canada specifies that congratulatory messages to the Prime Minister must be retained for a period of one year. In the private sector, a typical schedule might specify retention periods for items ranging from telephone message pads (30 days) to annual financial statements (permanent). Interestingly enough, fear of litigation may urge information systems managers to both archive (Skupsky 1993) and purge (Grady 1996) information on a regular basis.

6. CONCLUSION

This paper has illustrated some aspects of the relationship between social forgetfulness and information technologies. On the one hand, electronic information systems seem to almost naturally prevent a form of forgetfulness that might have been present to some extent in the paper-and-ink world. On the other hand, the opportunities offered by widespread collection of transaction-generated information seem to directly deny individuals the opportunity to shed their (digital) past.

Clearly then, the nature of public/institutional memory is dramatically changing due to the evolving character of information technologies. While preserving the opportunity for a second chance might have been easily achieved in the past, it has become increasingly difficult today. The ongoing balancing of "discard and forget" and "preserve and evaluate" has been skewed in favor of the latter. [BACKUP THIS CLAIM] Unless data retention issues are addressed explicitly as part of a comprehensive approach to personal privacy, we gradually move into a society displaying little if any social forgetfulness, little if any opportunity to move on beyond one's past and start afresh.

Robert Gellman (1998) warns that despite the fact that our three case studies exhibit a desire for some form of social forgetfulness, there is a trend, in all three cases, towards increasing maintenance of data: more juveniles are being tried as adults; bankruptcy law is being tightened (Johnston 1997); and limitations for data retention in credit reporting is being undermined by other, non-regulated, information services. Thus, the policies we use as examples are being undermined both by technology and the limitations of current law. There is a dual movement of technology overwhelming all attempts to control it, and of old limits fading. Thus, there is a need for a *reaffirmation* of the social value of forgetfulness.

There are many interesting ways in which we could further this initial exploration. For example, it would seem both interesting and necessary to clarify the relationship between social forgetfulness and privacy. Questions over social forgetfulness are usually raised within considerations of privacy and confidentiality: is this adequate, reductionist, or merely convenient? Privacy is sometimes decried as a catch-all category for a number of distinct social values, a phenomenon which may contribute to the conceptual muddleness of the privacy concept. Is this the case with regard to social forgetfulness?

We could also wonder, as Oscar Gandy has, about the obscure discrepancy between the rights enjoyed by corporations as legal personas, and those of individuals:

> "Corporations, unlike individuals, can be rather easily dissolved and formed anew on action of their boards of directors. Why should corporations as fictional persons already have rights that natural persons still long to enjoy?" (Gandy 1993, p. 225)

On what arguments have corporations been afforded a right that exceeds, both in its scope and ease of access, anything currently available to individuals? This is interesting!

We have pointed to a number of areas where society has granted individuals (limited) rights to social forgetfulness. What seems needed at this point is both an explicit reaffirmation of this value within existing privacy policies, and a comprehensive approach to information retention in the light of new modes of data collection and corresponding changes in the nature of social memory.

REFERENCES

Agre, P. (1997) Beyond the Mirrord World: Privacy and the Representational Practices of Computing, in *Technology and Privacy: The New Landscape* (P. Agre and M. Rotenberg, eds.). Cambridge, Mass.: The MIT Press.

American Press (1997) Soupçons sur la confidentialité des téléphones portatifs, *Le Devoir* (Montréal), December 29, A-8.

Associated Credit Bureaus (1975) How to comply with the Fair Credit Reporting Act, in *Consumer Credit Compliance Manual*. Rochester, New York: The Lawyers Co-operative Publishing Co.

Bean, P. (1981) Punishment: a philosophical and criminological enquiry. Oxford: Martin Robertson.

Charmichael, David W. (1998) Personal communication, Dec. 10, 1998.

Chaum, D. (1992) Achieving electronic privacy. *Scientific American* **267**(2): 96-101.

Chaum, D. (1985) Security without identification: transactions systems to make big brother obsolete, *Communications of the ACM* **28**: 1030-1044.

Chaum, D.(1981) Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* **24**: 84-88.

Chaum, D. and Evertse, J.-H. (1987) A secure and privacy-protecting protocols for transmitting personal information between organizations, in *Advances in Cryptology —CRYPTO '86 Proceedings* (A. M. Odlyzko, ed.). Berlin: Springer-Verlag.

Clarke, R. (1994) Human identification in information systems: Management challenges and public policy issus. *Information Technology and People* **7**(4): 6-37.

Davies, S. G. (1997) Re-engineering the right to privacy, in *Technology and Privacy: The New Landscape* (P. Agre and M. Rotenberg, eds.). Cambridge, Mass.: The MIT Press.

Douglas, M. (1980) How institutions think. Syracuse, New York: Syracuse University Press.

European Council (1981) Convention for the protection of individuals with regard to automatic processing of personal data, *European T. S.* no. 108.

[European Directive] (1995) Directive 95/46/EC of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Communities*, November 23, 1995, L. 281 p. 31.

Flaherty, D. (1989) Protecting privacy in surveillance societies: The Federal Republic of Germany, Sweden, France, Canada, and the United States. Chapel Hill, NC: The University of North Carolina Press.

Foucault, M. (1975) Surveiller et punir: naissance de la prison. Paris: Gallimard.

Gandy, O. H. Jr. (1993) The Panoptic sort: A political economy of personal information. Boulder, CO: Westview Press.

Gellman, R. (1997) Does privacy law work?, in *Technology and Privacy: The New Landscape* (P. Agre and M. Rotenberg, eds.). Cambridge, Mass.: The MIT Press.

Gellman, R. (1998) Personal communication, (date).

Guarino-Ghezzi, S. and Loughran, E. (1996) Balancing juvenile justice. New Brunswick, NJ: Transaction Publishers.

Grady, P. R. (1996) Discovery of computer stored documents and computer-based litigation support systems: Why give up more than necessary, *The John Marshall Journal of Computer & Information Law* **14**: 523-553.

Hacking, I. (1990) The taming of chance. Cambridge: Cambridge University Press.

IKG (1997) Federal Act Establishing the General Conditions for Information and Communication Services, 13 june 1997.

Johnston, D. C. (1997) Narrowing the bankruptcy escape hatch. *The New York Times*, October 4, B-9.

Lapierre, N. (1995) Changer de nom. Paris: Stock.

Mack, J. (1909) The juvenile court. *Harvard Law Review* **23**.

Marx, G. T. (1986) The iron fist and the velvet glove: Totalitarian potential within democratic structures, in *The Social Fabric: dimensions and issues,* (James F. Short, ed.). Beverly Hills, CA: Sage Publications.

Marx, G. T. (1988) Undercover: Police surveillance in America. Berkeley: University of California Press.

Mattison, R. (1996) Data warehousing: Strategies, technologies, and techniques. New York: McGraw Hill.

McNamara, R. M., Jr. (1973) The Fair Credit Reporting Act: A legislative overview. *Journal of Public Law* **22**: 67-101.

Nietzsche, F. (1997) Untimely meditations. Cambridge, New York: Cambridge University Press.

Nora, P., ed. (1984-1992) Les lieux de mémoire. Paris: Gallimard (in English, Realms of memory: Rethinking the French past. New York: Columbia University Press).

O.E.C.D. (1980) Recommendation of the Council concerning guidelines governing the protection of privacy and transborder flows of personal data. Doc. C(80) 58 final.

Parsloe, P. (1978) Juvenile justice in Britain and the United States: The balance of needs and rights. London: Routledge & Kegan Paul.

Regan, P. (1995) Legislating privacy. Chapell Hill, NC: University of North Carolina Press.

Reiman, J. H. (1995) Driving to the panopticon: a philosophical exploration of the risks to privacy posed by the highway technology of the future. *Computer & High Technology Law Journal* **11**: 27-44.

Rule, J. B. (1973) Private lives and public surveillance. London: Allen Lane.

Schwartz, P. M., and Reidenberg, J. (1996) Data privacy law: A study of United States data protection. Charlottesville, VA: Michie.

Sheldon, J., ed. (1994) Fair Credit Reporting Act. National Consumer Law Center, Consumer Credit and Sales Legal Practice Series.

Skupsky, D. S. (1993) Establishing retention periods for electronic records. *Records Management Quaterly* **27**(2): 40, 43-43, 49.

Sullivan, T. A., Warren, E., and Westbrook, J. L. (1989) As we forgive our debtors: Bankruptcy and consumer credit in America. New York: Oxford University Press.

Sweeney, L. (1997) Weaving technology and policy together to maintain confidentiality. *Journal of Law, Medicine and Ethics* **25**: 98-110.

Thomas, R. (1998) Police switch on the candid camera that knows your face. The Observer, 11 October 1998

Turner, F. J. (1986 [1920]) The frontier in American history. Tucson, AZ: The University of Arizona Press.

Vedder, A. H., Schreuders, E. and Van Kralingen, R. (1998) Knowledge discovery in databases and de-individualization. Paper presented at the "Computer Ethics: Philosophical Enquiry (CEPE'98)" conference, London School of Economics, December 13-14, 1998.

Virginia State Crime Commission (1996) Final report of the Virginia State Crime Commission on juvenile records retention study to the Governor and the General Assembly of Virginia. House document no. 38, Richmond: Commonwealth of Virginia.

Wacks, R. (1989) Personal information: Privacy and the law. Oxford: Clarendon Press.

Warren, C. (1935) Bankruptcy in United States history. Cambridge, Mass.: Harvard University Press.

Westin, A. F. and Baker, M. A. (1972) Databanks in a free society: Computers, record-keeping, and privacy. New York: Quadrangle/New York Times Book Company.

Willier, W. F. (1971) The Fair Credit Reporting Act: What is an attorney to do. Brighton, Mass.: National Consumer Law Center, Boston College Law School.