Abstract:
Most contemporary conceptions of privacy are based on a notion of a separation between the individual and the environment. On the one side of this boundary lies “the private” on the other lies “the public”. The struggle over the protection of privacy is about defending this boundary and the individual’s ability to determine who can access the private under what conditions.

Such a conception of privacy, far from being universal, is in fact historically specific. Its rise (and decline) is part of a particular cultural condition connected to the dominance of print media from the 16th to the 20th century. As electronic communications rise in importance, print culture, part of which is the notion of privacy, erodes.

Reacting to this development, academics and the general public have been concerned with the preservation of privacy for more 35 years, precisely the period when the shift from print to electronic communication accelerated. During most of this time, focus of privacy advocates was on the development of policy frameworks. Despite successes, this strategy has, by and large, failed to stop the erosion of privacy.

More recently, a new approach to privacy protection which promises to restore personal privacy has been developed in the context of the Internet: Privacy Enhancing Technologies (PETs).

This article analyses some of the key PETs and concludes that they, too, fail to protect privacy in the electronic environment. This supports the thesis of the historical situatedness of privacy and raises troubling questions for privacy advocacy in the long term.

1. Introduction

While a stringent definition of privacy is elusive, it is safe to say that at the core of our concepts of privacy is the notion of a separation between the
individual and the environment. On the one side of this gap is “the private” and on the other “the public.”\(^1\) The boundary, or gap, between the two spheres is controlled by the individual. Invasion of privacy means the intrusion of a “public” actor into the realm of the “private” without the individual’s consent. This notion is clearly visible in the approach to privacy prevalent in Europe, particularly in Germany. Here, privacy is translated into “informational self-determination”, that is, the right of individuals to police the boundary, to decide themselves which information they are willing to disclose under what conditions.\(^2\)

Such a conception of privacy, as I will argue, is far from universal, rather it is historically specific. It is part of a particular cultural condition connected to the dominance of print media from the 16\(^{th}\) to the 20\(^{th}\) century. As electronic media begin to supersede print media as a dominant mode of communication in our society, our notions of privacy, so closely associated with print culture, are becoming problematic.

It is not surprising, then, that over the last 35 years, the issue of personal privacy has been continuously on the agenda of academic research as well as the general public.\(^3\) For a long time, the struggle for privacy protection has focussed on the development of a suitable policy framework. Central to this project was the development of the *Fair Information Principles* by the Organization for Economic Cooperation and Development (OECD) in 1981,\(^4\) providing much of the conceptual basis for legislation in Europe and North America, culminating in the European Union’s Directive on Data Protection in 1995\(^5\) and the corresponding Canadian privacy legislation (Bill C-6) in 1999. In the US, privacy legislation has remained comparatively haphazard.\(^6\)

Despite such significant advances on the policy front, recent studies have argued that surveillance – the collection and compilation of personal information – is becoming an ever more central feature of contemporary societies and, consequently, privacy is continuing to erode. The sphere of the
private to which the individual can control access is shrinking, particularly on an informational level.⁷

In part as a reaction to the insufficiency of legislation, in part as a response to more general technological developments (the emergence of the Internet as mass medium) a new strategy has emerged which is making promises to halt, if not reverse, this trend: Privacy Enhancing Technologies (PETs). They aim to furnish individuals with the necessary tools to create their own privacy in the electronic environment.

This article investigates aspects of the voiding of privacy in two steps. First is a historical sketch outlining the connection of privacy to the rise and decline of print culture. The historical section leads into an empirical section examining the most advanced, and in many ways most promising, strategy to safeguard privacy in the electronic environment: Privacy Enhancing Technologies (PETs). The results of my analysis are overwhelmingly negative. PETs, despite their technological sophistication, cannot protect privacy. This gives credence to the argument that the notion of privacy is inapplicable, or at least very problematic, in an electronic environment. In the conclusion of this article, I will touch upon troubling questions this raises about the long-term salience of current privacy advocacy.

2. Privacy in historical perspective

As modern Westerners, we tend to think of the separation of the “private” and the “public” as quasi-natural.⁸ However, rather than being universal, the origins of this notion of a gap are historically situated. One of the first to articulate it was the writer Michel de Montaigne (1533-1592). For him, the human experience was comprised of two worlds – the interiority of the self and the exteriority of the world. To illustrate the relationship between the two domains, he used the metaphor of a house. Human beings, he wrote, have a front room that faces the street where they meet and interact with others, but they always needed to be able to retreat into a back room (their
most private self) where they can reaffirm the freedom and strength of their intimate identity and reflect upon the vagaries of experience. He wrote: “We must reserve a little back-shop, all our own, entirely free, wherein to establish our true liberty.”

Today, this concept of the difference between the inner self and the outer world has become so pervasive that it is hard for us to appreciate how counterintuitive this idea must have seemed to most of Montaigne’s contemporaries. Liberty, he argued, was to be established not by acting in the world, but by separating from the world. This separation was not to serve religious contemplation, as it had before; on the contrary, it constituted an attempt to better understand one’s position in the world. Liberty was based on an individual’s right and ability to have an opinion, and this opinion could only be formed in private. Only by distinguishing the public (front room) from the private (back room), was it possible to carve out the space in which this type of freedom could be achieved.

The underlying distinction that separates the front from the back room, however, is not fixed or uni-dimensional. As Gary Marx argues,

> public and private ... are best conceptualized as multidimensional (with dimensions sometimes overlapping or blurred and at other times cross cutting or oppositional), continuous and relative, fluid and situational or contextual, whose meaning lies in how they are interpreted and framed.

On the level of communication, the gap between the two domains can be found in the separation between the knower and the known. Publishing a book, for example, enables the author to communicate his knowledge without revealing himself in the process. This is most evident in cases of anonymous publishing. “Thomas Pynchon” is world famous for his extraordinary novels, but as an individual person, he has been able to remain unknown. Even if an author’s real name is printed on the book jacket, for most readers he or she will essentially remain unknown and unreachable.
Hence reading the book of a dead author is essentially not different from reading the book of a living author. The same separation also structures the reader’s experience. Books, and other printed materials, enable us to learn about the world without revealing anything about ourselves. Unless readers choose to identify themselves, an author has very little knowledge of the book’s audience. This, of course, is different online. Each viewing of a page is logged, and there are numerous ways in which it is possible to follow the reader around.

Over the last five decades, many communications scholars have argued that it is precisely the emergence of print culture, based on the spread of the printing press, that gave rise to our notions of individuality and privacy.\textsuperscript{11}

In an oral (or manuscript) culture, sustained thought was dependent on direct communication – the explicit, real time, face-to-face interaction of at least two parties. Hence, opinions and thoughts could never be truly private, thinking was not a solitary process. This changed with the availability of cheap printed materials. As Walter Ong notes:

\begin{quote}
Print was also a major factor in the development of the sense of personal privacy that marks modern society. It produced books smaller and more portable than those common in a manuscript culture, setting the stage psychologically for solo reading in a quiet corner, and eventually, for completely quiet reading. In a manuscript culture ... reading had tended to be a social activity, one person reading to others in a group.\textsuperscript{12}
\end{quote}

From this point of view, then, it is not surprising that it was Montaigne who developed the modern notion of the importance of the private. He was one of the first authors who made extensive use of a large, personal, secular library. This library was so important that he dedicated an entire tower of his castle to it and he spent much of his time there, comparing notes from books and contemplating about the different cultures and experiences contained in them.\textsuperscript{13} While clearly expressing a modern sensibility, Montaigne was a transitional figure between a world in which no notion of the private existed
and one in which it was to be taken for granted. For Montaigne, the “back room” was not yet something that existed naturally, but something that still needed to be argued for so that it would be established. Two centuries later, these notions are considered universal norms and inalienable human rights which provide the foundation for the modern political order.\textsuperscript{14}

The notion of privacy, one could argue, was an unintended consequence of the emergence of a new mode of communication: print. Privacy can be understood as part of the specific culture that the printed word’s dominance created, because print favours one-way over two-way communication. The author talks through the book, the reader reads the book alone. The author reveals without being revealed, the reader learns without being learned about.

While the separation between the private and the public was never without its own set of contentions, print’s physical nature has ensured that the gap between the two domains was maintained fairly reliable and unproblematic. There was simply no efficient way for authors to observe readers, even if they wanted to. As print culture became more deeply entrenched in Western culture, privacy, its unintended effect, became seen as one of society’s central virtues.

Hand in hand with the spread of electronic communication, the gap that separated private from public was bridged in unexpected ways, and the two spheres began to intrude upon each other. As Gary Marx notes,

\texttt{where does (and should) the private person stop and the public person begin? These questions were relatively more settled before new technologies appeared that suddenly give meaning to here-to-fore meaningless (or unobtainable), and therefore inadvertently protected, personal information.}\textsuperscript{15}

While Marx talks about new technologies that measure “brainwaves, body radiation and pheromones,” he could just as well have spoken about the way the personal information of the reader of a printed work is inadvertently
protected, in contrast to the highly observable reader of online publications. Or he could have talked about the way coins and bills inadvertently protect the privacy of buyer and seller. Perhaps, at least in the case of books and cash, it would be more accurate to say that they inadvertently created privacy, since the very notion didn’t exist up to that point.

There is a symmetry here. One can argue that the concept of privacy emerged as part of a much larger socio-technological shift in our society’s main means of communication – from a bi-directional oral mode to a one-directional printed mode. Following that logic, the current shift to electronic communication, which is also bi-directional, is likely to make the notion of privacy in its traditional sense void. In such a society, it will become increasingly difficult to reveal without being revealed, and to learn without being learned about.

There are, of course, very significant differences between oral and electronic cultures. However, what they have in common is that the experience of the self in the world is less characterized by the presumption of an ontological separation between the interior and the exterior, and more by interdependence and relationality. As a part of this shift, privacy as it was based on print culture is eroding and our conceptions of it are increasingly problematic in the Network Society.

3. Privacy Enhancing Technologies (PETs)
As already mentioned, several decades of legislative lobbying have yielded an impressive number of laws and regulations aimed at the protection of privacy, but in the eyes of most observers, these haven’t stopped its erosion. As a response to the deficits of legislation as well as to general changes in the technological set-up of our societies, a new approach to the protection of personal privacy has emerged in the last couple of years. A new breed of technologies, so-called Privacy Enhancing Technologies (PETs),
have been developed to help individual users control the amount of personal information they disclose in an on-line transaction. These technologies promise to enable individuals to take control over how their data is being collected. The goal is to restore the balance of power between the individual who wants to retain privacy and many actors in the online environment who want to gather personal information. The ultimate goal of PETs is to make informational self-determination a practical reality.

PETs, then, represent a paradigm shift in the struggle for personal privacy. Policy frameworks – legislation and self-regulation – aim at minimizing the occasions in which violations of privacy are attempted by restricting certain practices. Their ideal is a situation in which the individual’s privacy is protected by default and individual acts of transgression can be dealt with through the policy framework.

The assumption underlying PETs is that this vision has failed almost completely, particularly online. Rather than being protected by default, surveillance is the default, invasions of privacy are attempted continuously and routinely. Indeed, the almost complete lack of privacy is the backdrop that motivates PETs and propels their development.

PETs, ideally, allow individuals to take action to protect their own privacy against frequent and unknown attempts to infringe upon it. Rather than relying on the state – or some industry association – to deal with the problem on a collective level, PETs are technologies designed to support action by the individual for the individual. PETs recognize that electronic communications have massively increased scope of surveillance and aim to remedy this situation on the same technological level: a technological fix for a technological problem.

These technologies can be classified into three categories:21

- privacy through proxy
privacy through informed consent
privacy through untraceability

Not included in this classification, and my review, are encryption technologies, such as Pretty Good Privacy (PGP),\textsuperscript{22} a popular application to encrypt email messages, or the Secure Socket Layer (SSL), the standard protocol to secure the transfer of sensitive data (for example, credit card numbers) over web browsers. The ability to protect sensitive information from unauthorized access is primarily an issue of security. Even though security and privacy are often conflated in the general discussion, there is no necessary positive relationship between the two. A secure system does not need to be private.\textsuperscript{23}

**Proxy-Based Privacy: Remailers, Freedom™, and Others**

Privacy through proxy is perhaps the most frequent approach taken when building PETs. The basic idea is that somewhere within the chain of communication links connecting two parties, a specialized entity is inserted that blocks the receiver from seeing the sender. The simplest embodiment of this idea is the “Type I,” or “cypherpunk” remailer. It simply strips off the sender’s information from the header of the email before forwarding it to the recipient. In the ‘from:’ field of the email, the remailer’s address appears, often with a remark that the remailer is not the original sender. If several of these remailers are joined into a chain, it becomes virtually impossible to unmask the original sender.\textsuperscript{24}

The disadvantage of this approach results in what one could call the “anonymity paradox.” The only way to have a sustained anonymous communication on the Internet is to communicate in a very public forum that is small enough so all parties can see the flow of messages, but large enough so that membership in the forum does not offer a clue to the identity of the anonymous contributors. The cypherpunk email list was such a forum in
which, occasionally, anonymous discussions took place. However, the cypherpunks list was very unusual, because most of its members were deeply, if not fanatically, devoted to issues of privacy and anonymity. More often, anonymous communication is one-way: a dissenter publishes secret information, or an aggressor sends a death threat.

A more sophisticated form of email privacy via proxy is the pseudonymous remailer. Such remailers replace the original sender’s information with a pseudonym. However, the remailer retains the original email address, which makes it possible to reply to the anonymous email – the remailer forwards it to the original sender. Such remailers are similar to anonymous P.O. boxes. The advantage of this method is that it is possible to have pseudonymous email exchanges directly between individuals. However, the disadvantage is that the remailer itself is an obvious point of attack. This was dramatically demonstrated for the first time in February 1995, when the Finnish police raided anon.penent.fi, one of the most popular remailers, with 200,000 accounts at the time, run by Johan Helsingius.\(^{25}\) The Church of Scientology alleged that this remailer was being used to publish copyright-protected information, and Helsingius was eventually forced to reveal the identity of at least one user. The following year, he shut down his service.

Perhaps with this incident in mind, Zeroknowledge, a Montreal-based company, developed a software product called Freedom™. Freedom™ was a highly sophisticated technology that allowed the user to create up to five pseudonyms, which she could use in different contexts for web browsing and email. The sophistication of the system was based on the fact that not even Zeroknowledge could match the pseudonym with the real address, removing the kind of vulnerability suffered by Helsingius. Freedom™ was launched in 1999 to rave reviews. The privacy community was impressed by the technical strength of the product, and Ann Cavoukian, Ontario's Information and Privacy Commissioner, called Zeroknowledge the “Mercedes-Benz of anonymizer-technology companies.”\(^{26}\) During the height of the Internet...
boom, Zeroknowledge was awash in venture capital and distributed T-Shirts with the slogans such as “Internet Freedom Fighter” and “Privacy is Sacred.” However, despite broad acclaim from reviewers, Freedom™ did not survive the climate change after the burst of the dot.com bubble. In October 2001, the service was shut down. Even though no numbers were released, it was widely understood that a combination of low subscribers and a change in business strategy made it commercially unviable for Zeroknowledge to continue to support its former flagship product.

Safeweb.com, a late but prominent entry into the arena, experienced similar difficulties in the marketplace. Safeweb provided a web-based proxy that users could access to browse the Internet without the web site’s operators being able to track them. Contrary to Zeroknowledge, Safeweb was not extremely sophisticated and, being a simple proxy, it had the same vulnerability as Helsingius’ system. However, it focussed on ease-of-use, and created a simple point-and-click interface that required no technical knowledge or registration.

From an economic point of view, the greatest weakness of the privacy-through-proxy approach is the fact that they are resource-intensive services which scale poorly. They effectively channel all users through a central hub – the proxy server. This introduces a bottleneck which demands a lot of computing power and bandwidth from the central proxy, both of which are expensive. Furthermore, the costs of running a proxy linearly increase with its number of users. In other words, the more popular a service becomes, the more expensive it becomes for the provider. This seriously limits the capacities of grassroots services and forces commercial ones to depend on user charges, particularly since paid advertisements are only a limited fundraising option. Safeweb was forced to acknowledge this problem in late 2001. It closed its public service and shifted its focus to “enterprise applications,” like Zeroknowledge, hoping to find a clientele more willing to pay for a PET.
Anonymizer.com, one of the oldest and most prominent commercial providers of PETs, is still in existence. However, despite its excellent brand name and wide recognition, it has not been able to sign up more than 17,000 paying subscribers, making it an established but very small player.\textsuperscript{27}

The prospects for privacy-through-proxy are not very rosy. If the experience of the last 5 years is any indication, individual users are unwilling to pay for such services. Consequently, many commercial providers have shifted their target audience from consumers to companies. Free grassroots services are still quite numerous, but are severely restricted in the amount of volume they can handle. Though they play an important role for a small number of users on a day-to-day basis, or for average users in exceptional circumstances, they cannot break out of their specialized niche for the scaling reasons mentioned above. This might change with new, decentralized, peer-to-peer approaches that avoid scaling issues and are therefore more appropriate for grassroots employment. Some new products have been announced, but they have yet to be widely employed.\textsuperscript{28}

**Privacy Through Informed Consent: P3P**

It is also possible for privacy to be achieved through a “fair negotiation” between the individual who is asked to release personal information and the institution that provides a service in exchange. If individuals gives informed consent to the exchange of data, then their privacy rights are adequately respected. This concept of privacy underlies Platform for Privacy Preferences Project (P3P).

Many online service providers have privacy policies, and some of these policies are even certified by more or less trusted third parties, such as TRUSTe.\textsuperscript{29} However, even those web sites with certified privacy policies exhibit quite a variety of approaches to handling personal data. What makes
things more complicated for users is that fact that these policies are often hard to find on the actual web site, and when found, they are often hard to understand. Most policies are written by lawyers whose mandate was to protect the owner of the site, rather than efficiently inform its casual users. Furthermore, these policies are often changed according to the strategic plans under which the sites operate. Often, such changes are made without prominent public notice. All of this makes the privacy policy a valuable but highly insufficient tool for protecting the users’ privacy interests. In practice, most users do not know if a site has a privacy policy at all, and even less often do they know the degree to which this privacy policy conforms to their own expectations.

To address this problem, the World Wide Web Consortium (W3C), a not-for-profit group designing open standards for the Web, has been developing the Platform for Privacy Preferences Project (P3P). P3P aims to be become the industry standard for encoding a web site’s privacy policy in a machine-readable form. As the W3C writes:

P3P is a standardized set of multiple-choice questions, covering all the major aspects of a Web site's privacy policies. P3P-enabled Web sites make this information available in a standard, machine-readable format. P3P-enabled browsers can "read" this snapshot automatically and compare it to the consumer's own set of privacy preferences. P3P enhances user control by putting privacy policies where users can find them, in a form users can understand, and, most importantly, enables users to act on what they see.

The P3P standard is designed for two purposes: to simply and automatically communicate a web site’s stated privacy policies to users, and to determine how they compare with the user’s encoded privacy preferences.

P3P does not set minimum standards for privacy, nor does it monitor whether sites adhere to their own stated policies. These issues lie outside the technical mandate that the W3C has set for itself. P3P is intended to
complement both legislative and self-regulatory programs that can help to set and enforce website policies. Self-regulation is promoted, with varying degrees of effectiveness, by organizations such as the aforementioned TRUSTe.

After close to three years in development, P3P has matured into version 1.0, launched in mid-2002. Microsoft has announced that it will include a P3P module in the next release of its Internet Explorer, the most popular web browser.

The privacy community’s reception of P3P has been mixed. Some people cautiously support it, while others see it as a flawed approach that does more damage than good. Among the former are the Center for Democracy and Technology, and the Office of the Information and Privacy Commissioner of Ontario. In a joint publication, they state: "As privacy advocates, we believe that – armed with more information – individuals will seek out companies that afford better privacy protection." In this view, P3P is seen as empowering users to choose companies with good privacy policies, which will translate into competitive pressure that forces companies to upgrade their policies.

This view is quite optimistic. It assumes a) there is competition that offers the same service with more privacy, and b) that the "opportunity costs" for finding this competing service are lower than the perceived gain in privacy. Both of these assumptions might not be true, as Karen Coyle, of the Computer Professionals for Social Responsibility (CPSR), has pointed out. The current state of copyright and IP laws ensure that certain types of content will be offered only by a limited number of outlets. For example, if I want to read the New York Times, only one provider can give me access to the information. Syndication of content might mitigate this, but the incentive to syndicate is lower online than it is offline. The main protocol for content syndication is Resource Description Framework (RDF), which syndicates
only links to the original story, rather than the content itself. Furthermore, as long as advertisement plays a dominant role in financing online content, most sites are under the same pressure to deliver customer data, and hence their privacy policies are likely to be similar. In addition, the actual opportunity costs of finding alternative services might be higher than a popular Internet myth – everything is one click away – suggests.

Increasingly, Internet traffic has become highly centralized, with a small number of web sites attracting a disproportionately high percentage of overall traffic. This suggests that many users are not aware of, or at least do not use, a broad range of services and information sources beyond well-established mega-brands. For them, finding alternatives is particularly difficult and time-consuming.

Some critics within the privacy community raise deeper issues. Karen Coyle’s radical position is that P3P represents a “tacit acceptance of the great increase in the tracking and monitoring of our minor activities that takes place over the Web.”

In this view, rather than fighting the trend towards monitoring users, P3P makes it manageable. Rather than protecting personal information, P3P is a tool to negotiate its transfer.

Perhaps the most problematic aspect of P3P is that it isolates privacy from the other dimensions that make up an interaction. Only the least important interactions, likely to involve only limited personal information, will be broken just because of the privacy policy. For most people, privacy concerns are balanced, at various weights, against other concerns that motivate action. If a user, even a well-informed user, must constantly chose whether or not a privacy policy is sufficient – even if it is only a simple yes/no click – then the system runs the danger of breaking down under its own load. The problem of managing cookies through web browsers is suggestive here. Most web browsers have the option to prompt users when a web site tries to plant a cookie on their hard drive. However, as many sites are using cookies, the prompts become so frequent, and so annoying, that this option is virtually
unlikely. All but the most zealous users will turn this function off sooner or later, particularly since banning all cookies is not a very practical alternative. Paradoxically, the more alerts such systems generate, the less valuable they become – a classic problem in the field of security technologies.39

In effect, P3P could end up forcing users to lower their own stated privacy expectations simply to reduce the frequency of alerts created by P3P-enabled browsers. This is particularly problematic if the browser sends out an alert for every site that is not P3P-enabled. At least initially, this will include the large majority of sites. However, if P3P warns only of adequately coded but insufficient policies, there is no real incentive to employ P3P on a web site. This is a well known problem of any voluntary rating system. But even if P3P becomes a widely adopted standard, the optimistic argument put forward by P3P proponents – that the power of user expectations will force web site operators to upgrade their privacy policy – remains questionable. P3P usage might well turn into a pessimistic scenario. The power of web site operators may well force users to reduce the sensitivity of their privacy modules in order to be able to function conveniently on the web. This would be particularly likely if the surrounding (self)regulatory framework does not already mandate minimum privacy standards.

As critics point out, P3P could hamper the creation of such a framework. It could serve as an excuse for industry and the (US) government, who do not want to see strong privacy protection, to further stall effective legislation in this area.40 Without legislation, P3P shifts the burden to the individual user, who must choose between protecting his/her privacy or accessing a particular web site. This is unlikely to contribute to reducing the incidence of online behaviour being monitored. However, with legislation, P3P could be a powerful tool to make this monitoring become transparent, as well as to identify which companies are compliant with the legislation and which are not.41 With binding legislation, the number of non-compliant web sites would likely be small, hence the alerts would be of real value.
Privacy as Untraceability: Freenet

Privacy, particularly in relation to freedom of speech, is often thought of as anonymity. Anonymity thought to be necessary so that the speaker can protect herself against prosecution initiated by those parties who do not want certain information to be released. The classic points in case are political dissidents in totalitarian systems and corporate whistle blowers. A free society requires the possibility to speak anonymously. This kind of privacy is what Freenet aims to enhance.

Freenet is not related in any way to the Freenets, pioneering community networking projects of the 1980s and early ’90s that provided people with free or cheap access to the Internet. Rather, it grew out of a research project launched in 1997 by Ian Clarke at the University of Edinburgh’s Division of Informatics. The initial version of the Freenet code was posted online as an open source project in the summer of 1999. Two and a half years later, the project has progressed to version 0.4. In other words, it is still in development mode and not yet considered ready for large-scale implementation. However, its main features have become clear.

Freenet’s main objectives are to provide an infrastructure that incorporates:

- anonymity for both producers and consumers of information
- deniability for storers of information
- resistance to third parties’ attempts to deny access to information
- efficient, dynamic storage and routing of information
- decentralization of all network functions

The World Wide Web does not achieve any of these objectives. Each resource is identified through a URL (Uniform Resource Locator), hence it is easy to pinpoint its location and, at the very least, the owner of the server on which it is stored. This provides a shortcut to censoring information on the Internet.
Today, most countries have legislation that makes service providers liable for knowingly hosting objectionable content. Since most service providers have no stake in the content, they often comply with simple requests for removal, even without a court order. Furthermore, each server maintains a log of the IP addresses, date and time that a given document was requested, making it possible to locate most users. Since most information is stored only on a few servers, it’s possible, though in practice not easy, to remove information. Since content is stored only in one place (and, perhaps, a few mirror sites) distribution is not particularly efficient. Sudden high demand can overwhelm and effectively bring down smaller servers. “Flashfloods” are such a regular occurrence on the web that they have their own term in Internet jargon: the slashdot-effect.\textsuperscript{44} Finally, there are several instances of centralized control on the Internet. Most importantly, the Domain Name System (DNS) translates easy-to-use domain names into computer readable addressing numbers. This creates an obvious point of control, which is currently managed by the Internet Corporation for Assigned Names and Numbers (ICANN).

To avoid what are seen as weaknesses in the current WWW model, Freenet employs a very different architecture. Instead of a client-server architecture, it uses a peer-to-peer network of nodes. In a joint paper, some of the main authors of Freenet summarize it as follows:

\begin{quote}
Freenet is implemented as an adaptive peer-to-peer network of nodes that query one another to store and retrieve data files, which are named by location-independent keys. Each node maintains its own local datastore which it makes available to the network for reading and writing, as well as a dynamic routing table containing addresses of other nodes and the keys that they are thought to hold.\textsuperscript{45}
\end{quote}

Central to the Freenet concept is the idea of “distributed caching”. A cache is a temporary store of data, and every web browser has one. In the Freenet model, every node stores transient data. This works because when a document is passed back along a chain of Freenet clients from where it is stored to the original requester, each client in the chain keeps a copy of the...
document. In order to avoid an infinite duplication of content, each client has an expiry mechanism that deletes data if nobody has requested it for a given period of time. In other words, content that is frequently requested is distributed throughout the network, while content that is never requested disappears. This architecture has the following properties:

- It creates the transience required to meet the goals of anonymity and persistence. There is no relationship between where a document is temporarily stored and where it originates. Because popular documents are multiplied, removing them from the network is extremely difficult.

- It lets small sites distribute large, popular documents without suffering from the slashdot-effect. The availability of information grows in proportion to the demand for it: if many people want that information, it will be stored multiple times across the network. With storage generally being cheaper than bandwidth, this is a more efficient system.

- This replication brings data close to those who want it. As in all peer-to-peer systems, "closeness" in Freenet has no geographical meaning. It refers only to the number of hops between the person who seeks the information and the node where the document is stored. The first request from node A to node E may have to pass through node B, C, D. The second and subsequent requests can be satisfied by node A directly. This not only improves the network’s efficiency, but also offers legal protection to the node’s owner through "plausible deniability" – even the person who actually requested the document can always claim that she too was merely a link in the chain of requests.

The second important aspect of Freenet’s architecture is that all content is encrypted. The host of a Freenet node cannot know what is on her server, because the documents are not only transient, but also unreadable without
decryption. In other words, the provider of a server cannot be made liable for knowingly distributing objectionable content.

Content is identified by a key, not by a location. For this to work, each node must have a table of keys that identifies locally-stored content, as well as the key tables of nodes “close” by. Since the content is unreadable until it is decrypted, full text searches are not possible, only key searches. In order to find information, a user needs to know the exact key that relates to it.

Freenet is not yet ready for wide implementation, and there are still significant technical problems that must be solved. Currently, the software is written in Java, which adds to its cross-platform compatibility, but is rather difficult to install and use, even for people with significant technical skills. Furthermore, the fact that one needs to know exact keys to find information significantly limits the network’s ability to find new content. Such shortcomings are not surprising, given that the project is barely three years old and has currently about 400 contributing programmers.

However, even if the technical problems are solved as currently planned, Freenet’s basic architecture introduces other problematic aspects. Since data is cached dynamically throughout the network, it is difficult for individual node owners to control the bandwidth that the system uses. Bandwidth, however, is one of the prime determining factors in pricing an Internet connection, making it difficult for node owners to control costs. This, combined with the fact that the owners have no control over the content stored on nodes, leads to the not very enticing situation that the owner might end up paying significant bandwidth costs for the transfer of unknown content. This is compounded by the fact that in order to protect the anonymity of information seekers, the requested documents pass along the same change of nodes as the request, rather than directly back to the querent. An advantage of this kind of information exchange, from a system perspective, is that each node only knows the node closest to it – not the
beginning or end of the chain – but there is the disadvantage that, from each individual node’s perspective, bandwidth usage is multiplied.\(^{48}\)

Distributed, dynamic caching has other disadvantages. In order to avoid the infinite multiplication of content, a mechanism for deleting extraneous information is necessary. Since node owners do not know what is on their nodes, this mechanism must to be automatic. Developers of Freenet are currently proposing that if a node becomes full, it then automatically deletes its least-requested documents, in order to make room for new files. In this way, popular material multiplies, and unpopular material disappears. Whether information is “unpopular” is determined by each node’s internal usage logs. In other words, on some nodes, certain files might be unpopular and disappear, while on other nodes they are preserved. Only when a file is not requested at all will it disappear. Freenet, nevertheless, is clearly not intended to be an archive, but a distribution channel for active content.

Another aspect of Freenet is that once a document is released onto the system, it cannot be retracted. This means also that it cannot be updated, corrected or otherwise improved. While clearly a plus from an anti-censorship point of view, this feature is problematic from most others.\(^{49}\) To justify this feature, Freenet developers operate with a very crude notion of censorship. Any blocking of information is censorship, no matter what the content, or veracity, of that information might be. For Clarke, there seems to be no difference between, say, political dissent and libel. As he writes: “Basically, you either have censorship, or you don't. There is no middle-ground.”\(^{50}\)

The most contentious aspect of Freenet at the moment relates to copyright and intellectual property. The system was designed with free speech in mind, but its architecture foils any attempt to restrict the free circulation of material, whether that material is a suppressed political manifesto, a leaked report, a copyrighted song, or illegal pornographic images. For the Freenet developers this is no problem, but a logical consequence of their
understanding of the tension between copyright and freedom of speech. As Clarke writes: “You cannot guarantee freedom of speech and enforce copyright law. It is for this reason that Freenet, a system designed to protect Freedom of Speech, must prevent enforcement of copyright.”

Freenet developers have a radical approach to building its privacy-enhancing infrastructure. Its slogan – “Rewiring the Internet” – is apt. Privacy, anonymity and free speech considerations are given paramount importance when considering how to survive in a hostile online environment, in which attacks can be technical as well as legal. Freenet is a (deliberately) crude system, which presents black and white social choices. Once released onto Freenet, information becomes very hard to control, for better or for worse. The only way to control what circulates on Freenet is to block the entire network, i.e. to make running a node itself illegal and/or to block all Freenet traffic to the extent that it can be distinguished from normal Internet traffic.

The system’s untraceability, then, comes at a high price. It forces users to release information that cannot be retracted. It also automatically destroys information that has not been requested for a certain period of time. Its inherent social crudeness will limit its applications, but not necessarily its viability or its social ramifications should it ever reach full-scale implementation.

4. Why PETs Fail to Enhance Privacy

The technologies reviewed above achieve their goals, or at least show potential to do so, only if those goals are defined very narrowly. Remailers make it impossible to identify the sender of an email, web-proxies make casual tracking of browsing patterns difficult. P3P, should it ever be implemented widely, gives users better information about a site’s privacy policy, and Freenet, at least conceptually, makes impossible to locate where Internet content is stored and thus protects speakers’ anonymity as well as
the free flow of information. Freenet also protects the owner of the servers from liability for content stored on their machines.

However, if we consider the goals of PETs to be more broad – to enhance privacy on the Internet for the majority of users – these technologies cannot be judged other than as failures.

As Freedom™ and Safeweb illustrate, there is no consumer market for privacy. If the last five years are any indication, commercial projects trying to sign-up paying users fail entirely, or are restricted to small niches offering specialized services to marginal markets. At the corporate level, market for PETs is beginning to emerge in order to comply with new privacy regulations, but the primarily goal there is to minimize corporate liability, rather than to protect users’ interests.

The difficulties of creating a consumer market for PETs is reflected in a well-known paradox: the overwhelming majority of people, when asked in opinion polls, claim to be concerned about their privacy. However, in practice, most users do very little to protect it.

Many privacy advocates explain this contradiction by the average users’ lack of information. People, they argue, simply don’t know the degree to which their personal information is being collected by commercial or government agencies, and are unaware of the potentially harmful consequences of this practice. This assessment is probably not incorrect, but as an explanation it is highly unsatisfactory. Few social issues have enjoyed so much media attention, and most people tend to inform themselves about issues that really affect them. When the Mad Cow disease hit Europe, people changed their behaviour. Why not with privacy?

I will offer three hypotheses that attempt to explain this paradox, and will expand my claim that PETs cannot protect privacy for most people most of
the time. The first hypothesis centers around the disconnect between action and effect, the second highlights the contextuality of the exchange of personal information, and the third concerns the imbalance between the provider and the individual users of a service.

First, there is a significant disconnect between action and negative effect, while there is often a direct connection between action and positive effect. Usually, when I give personal information in exchange for a service – for example, access to the New York Times web site – the gratification is instant. I fill out the form, log on and get the information, all in a matter of seconds. On the other hand, their tracking and compiling of my reading habits can occur over several months, but might never affect me, at least not at a conscious level.

Such a disconnect is not unique to issues of privacy. Many environmental problems face the same situation. For example, the problem of acid rain was recognized long before people started to do anything about it. Why? Because, say, driving to the mall has a direct positive effect – access to shops – while the compounded negative effects of acid rain always seemed to be something that would occur in the distant future. From the perspective of individual car owners, the benefits of having a catalytic converter installed were intangible. The costs, on the other hand, were very real and immediate. Had the catalytic converter remained an expensive “premium service” for people who care particularly about the environment, it would never have been widely adopted. Certain problems of collective concern are extremely hard to solve through voluntary, individual action. In the case of PETs, this disconnect contributes to make paying for privacy protection a very unattractive market proposition.

However, the analogy between PETs and catalytic converters is limited. Emissions of sulfur dioxide and nitrogen oxide are a bad thing. There are no instances in which such pollution is desirable. Hence the solution of whether
or not to install a converter that uniformly reduces emissions has, apart from the costs, no negative side-effects for users. This, of course, is different from the creation and compilation of personal information. This brings us to the second hypothesis.

Creating personal relationships between an individual and corporate or government agencies is not necessarily a negative thing. For instance, I want my bank to know who I am, and I want my credit card company to keep track of all purchases connected to my credit card number. The list of situations in which we want our personal information to be known to third parties is not only long, and growing longer, but also the composition of this list itself is highly personal. Some might oppose, say, police surveillance as intrusive, others might applaud it as a welcome security measure.

There is no one-size-fits-all solution. Not too long ago, we had only had a few personalized relationships with organized entities, mainly government agencies. Back then, it would have been relatively easy to devise a comprehensive framework for handling personal data. Now that electronic communication is personalizing nearly all our interaction with institutions, the contexts in which that information is generated and compiled have become so heterogeneous that we would need a very complex and flexible framework and a large toolbox to deal with each instance adequately. Anonymity is an option for only few of them, and necessary for even fewer.

The exchange of personal information occurs extremely frequently online and, from an individual’s point of view, is highly textured. In some circumstances it is necessary to be identified, but in others it’s desirable, doesn’t seem to matter at all, is just one nuisance among many others, or is something we want to avoid at all costs. Given the frequency of the problem and the context-dependent nature of the solutions, it becomes clear that a simple proxy is a highly inadequate answer, which gives people very little of value, and hence offers a poor market proposition. However, a proliferation
of tools increases their cognitive and financial costs beyond the point where the tangible benefits would justify these costs for all but the most paranoid. This, again, impedes the emergence of a market for PETs.

This problem is made worse by the fact that while, for the individual, the desirability of exchanging personal information is highly textured, for the receiving institution, thanks to advances in information processing, it is not. On the contrary, it’s rather uniform: the more, the better. This imbalance lies at the heart of the problem of P3P.

P3P promises to empower individuals to make informed choices. However, the situations in which choice is exercised are so imbalanced that they can hardly be called fair. As mentioned above, often these choices come in the form of take-it-or-leave-it. For example, to read the NYT there is only one place to go on the web. Furthermore, service providers have a lot to gain from gathering as much information as possible. The individual user, on the other hand, has very little to lose, particularly if the transfer of personal information is negotiated separately in each instance. This problem is similar to that of skimming money in frequent transactions. For the person doing the skimming, illegally subtracting a fractional amount from each transaction adds up to a substantial sum over a large volume of transactions. For an individual user, who’s looking at each transaction separately, the amounts being skimmed are almost imperceptible.

This power imbalance gives the provider a much better bargaining position than an individual user whose main incentive is to access a particular service. However, the service provider has an immediate interest in gathering personal information, whether to provide a better service, or to gather data that can be aggregated into a valuable resource. By this, I do not mean to imply that the provider has bad intentions, simply that the intentions of the service provider are directly related to one of the central promises of
electronic media and the delivery of personalized services: to give you what you want when you want it.

Electronic information processing allows providers to narrow down a large set of possibilities down into a smaller number of events based on the specific relationship between an individual and the service provider. At least to some extent, this empowers users to make their own choices, but it also intensifies the need for corporations to create personal profiles. Amazon.com is a good example of this wide-spread phenomenon. The site has such a wide selection of books on sale that many appreciate its added feature, which recommends books to us based on our profiled purchasing habits. There is no more scanning the (virtual) shelves, the selection would be too overwhelming.

The provider’s desire to collect personal information is further increased due to the online environment’s volatility. The faster the environment changes, the more important detailed and real-time information becomes. Knowing your customers is not a fancy service ideal, it’s an imperative in a fast-paced marketplace that deals with highly customizable, unstable information products.

Providing the user with easy-to-read information about privacy policies does not address this imbalance in a significant way. While there will be the odd exception, users will effectively have to accept whatever policy is offered to them, if they need the service and cannot get it elsewhere. P3P’s “informed consent” does not amount to a fair choice, in the same way that turning-off the cookie filter on our browsers does not indicate our voluntary acceptance of all cookies.

Freenet recognized that balanced negotiations between the infrastructure’s provider and its users is not possible. In order to protect its users from profile tracking, the system must be built so as to make this impossible. However, Freenet also illustrates the price that must be paid to obtain such
an architecture. Tracking and control are not necessarily bad actions, and by making them impossible in all situations their positive aspects are also disabled. Freenet’s attraction is thus limited to the margins, a fringe culture that is not without its merits, but excludes practices that a lot of people might value. Faced with the choice of either being tracked on the World Wide Web, or losing control on Freenet, most people will choose the former most of the time. However, it’s important to reiterate that this does not invalidate the Freenet’s potential to undermine copyright or protect anonymous speech, though its potential as a tool to protect privacy in a wide range of circumstances is limited.

5. Conclusion

Privacy – a notion based on the enforcement of a boundary between two distinct domains on the individual level – is not universal but emerged in relation to print culture. Over the last 35 years, legislation has not been able to stop the continuous erosion of privacy, and PETs, more recently, have shown to be useful only in very narrow domains. This double failure is an indicator how poorly privacy can be translated from the print to the electronic environment. The failure of PETs, which were developed specifically to prevent invasions of privacy where they occur ever more frequently, online, shows this quite clearly – and we have barely begun to integrate the Internet into everyday life. There is little indication that the boundary between the private and the public is not blurring.

Does that mean that Sun Microsystem’s CEO Scott McNealy was right to proclaim: “You have no privacy, get over it!” No. This would be the wrong conclusion to draw from the argument developed here. A better conclusion is to acknowledge the need to redefine privacy.

The conventional notion of privacy has become unworkable in an environment constituted by a myriad of electronic connections. As many
observers have noted, increasingly our societies are organized as networks underpinned by digital information and communication technologies. In a network, however, the characteristics of each node are determined primarily by its connections, rather than its intrinsic properties, hence isolation is an desirable option only in very exceptional cases.

So rather than fight those connections privacy advocates have to reconceptualize what these connections do. Rather than seeing them as acts of individual transgression (x has invaded y’s privacy) it is necessary to see them part of a new landscape of social power. David Lyon has suggested repeatedly a need to shift attention away from the individual towards the structural, away from privacy and onto surveillance. Surveillance, however, is not simply the result of frequent invasions of privacy, but a symptom of something else, of a new mode of organizing social relationships. David Lyon calls this “social sorting.” Such social sorting is not necessarily negative, rather, it can be seen as a way of organizing hyper-mobile societies. Surveillance is, however, a technique of power. It gives those who collect and process personal data important resources to shape the destiny of those whose information they hold.

What is necessary, then, is to find new ways of holding those with (data) power, accountable to whose who are effected by this power. As a question of social organization, the individual level – on which the notion of privacy focusses – is not the right place to address it. We can see this in the fact that even though privacy advocates are politically relatively influential – many jurisdictions even have institutionalized privacy commissioners – for every privacy invasion they expose, an unknown number occur unnoticed. As such, the conventional notion of privacy is becoming a liability, rather than a powerful weapon, in the struggles of the network society to curb new and unchecked forms of power.
ENDNOTES:


2 For an overview of recent approaches, see Bennett, Colin; Grant, Rebecca (eds) (1999). *Visions of Privacy*. Toronto: University of Toronto Press


On why Montaigne was so concerned with the peculiarities of different cultures, Elisabeth Eisenstein writes: “He could see more books by spending a few months in his tower study than earlier scholars had seen after a lifetime of travel. When explaining why Montaigne perceived greater ‘diversity and conflict’ in the works he consulted than medieval commentators in an earlier age, something should be said about the increased number of texts he had at hand” (p.44). Eisenstein, Elisabeth, L. (1983). The Printing Revolution in Early Modern Europe. Cambridge, UK: Cambridge University Press.


Ong (1980, p.3) writes: “The electronic age is also an age of ‘secondary orality,’ the orality of telephones, radio, and television, which depends on writing and print for its existence.”

I will expand on this in the conclusion of this article.

See, Bennet (1992), and Flaherty (1989)


Increasingly, there are also PETs for institutions, however, in the following, I will focus on those aimed at individual users.


There is also a Type II remailer, or MixMaster, that fragments messages into fixed-size packets, which are then bounced to different remailers in the chain. This greatly decreases the feasibility of traffic analysis.


http://www.peek-a-booty.org

http://www.truste.org/

http://www.w3.org/P3P/ [13.12.01]
This narrow mandate is quite deliberate. “We do not want specification and standard settings bodies determining public policy. W3C does not wish to become the forum for public policy debates. We don't want to cede the development of substantive policy to technical organizations.”


Opportunity costs refers to what is being lost in order to gain something else, for example, the time spent finding the competing service.


Cookies are small files stored on the user’s hard disk that identify him/her vis-à-vis a web service. A cookie can store access passwords or information to customize a web site. It can also be used to track the user’s surfing patterns within and across web sites.


45 Hong (2001)

46 As a comparison, the Linux project is more than 10 years old, and builds on software released in the mid 1980s.


48 This might become less of a problem as bandwidth becomes cheaper.

49 As the Freenet FAQ explains: “Proposals for a more useful mechanism are being evaluated, and one of them will probably be implemented in an upcoming version of the protocol. For example, documents could optionally be inserted with public keys attached, and only updates signed by the corresponding private keys would be accepted. Unsigned documents would be immutable. Alternately, some type of versioning system could keep track of all previous versions of documents.” <http://freenet.sourceforge.net/index.php?page=faq> [14.12.2001]


51 Ibid.

52 This is ironic, given that many players in the PETs-field hold strong libertarian beliefs.

53 For an overview of recent privacy surveys, see http://www.privacyexchange.org/iss/surveys/surveys.html

54 Concepts such as “ubiquitous computing” envision virtually every object – cars, fridges, heating systems etc – connected to the Internet.


59 The most extreme negative case of the power of surveillance as social sorting is documented in Black, Edwin (2002). IBM and the Holocaust: The Strategic Alliance between Nazi Germany and America’s Most Powerful Corporation. New York: Three Rivers Press, Random House