The Unintended Consequences of Privacy Paternalism

Ann Cavoukian, Ph.D.
Information & Privacy Commissioner
Ontario, Canada

Dr. Alexander Dix, LL.M.
Commissioner for Data Protection and Freedom of Information
Berlin, Germany

Khaled El Emam, Ph.D.
Canada Research Chair in Electronic Health Information
University of Ottawa

March 5, 2014
Acknowledgements

The authors wish to acknowledge Michelle Chibba, Director, Policy & Special Projects, with a special thank you and appreciation to Fred Carter, Senior Policy & Technology Advisor. In addition, our thanks go to Catherine Thompson, Regulatory and Policy Advisor and Oren Weichenberg, Legal Counsel for their contributions to this paper.
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>I. Response to “Data Protection Principles for the 21st Century:</td>
<td></td>
</tr>
<tr>
<td>Revising the 1980 OECD Guidelines</td>
<td>5</td>
</tr>
<tr>
<td>Diminishing Consent Weakens Essential Privacy Protections and Freedoms</td>
<td>6</td>
</tr>
<tr>
<td>Removing Purpose and Use Limitation is Inversely Related to Accountability</td>
<td>7</td>
</tr>
<tr>
<td>Privacy Breeds Innovation — Not the Reverse!</td>
<td>10</td>
</tr>
<tr>
<td>More Accountability, Definitely! Relying on After-the-Fact Redress, Why?</td>
<td>11</td>
</tr>
<tr>
<td>No Consensus on “Harm” or Agreed-upon Ways to Measure or Mitigate “Harms”</td>
<td>11</td>
</tr>
<tr>
<td>Privacy by Design Innovates by Extending User Control</td>
<td>13</td>
</tr>
<tr>
<td>II. The Next Evolution of FIPPs: <em>Privacy by Design</em></td>
<td>13</td>
</tr>
<tr>
<td>Privacy by Design Enhances Accountability — Across the Board</td>
<td>14</td>
</tr>
<tr>
<td>Towards a Design-Aware Future</td>
<td>15</td>
</tr>
<tr>
<td>The Enormous Value of De-identification</td>
<td>15</td>
</tr>
<tr>
<td>III. Conclusion</td>
<td>18</td>
</tr>
<tr>
<td>Appendix</td>
<td>19</td>
</tr>
<tr>
<td>Bibliography</td>
<td>21</td>
</tr>
</tbody>
</table>
This paper sets out to reinforce the fundamental privacy principles of purpose specification and use limitation that prescribe limits to the collection and use of personal data. We respond to a recent proposal to dramatically revise the OECD Fair Information Practice Principles (FIPPs) in the era of Big Data, Cloud Computing and the Internet of Things. The co-authors of the proposal argue that the current practice of “Notice and Choice” is deeply flawed in today’s era of ubiquitous data availability, and that the principles of Purpose Specification, Collection Limitation and Use Limitation be diminished in favor of greater emphasis on ensuring accountability by data users/controllers. We believe the proposal reflects a paternalistic approach to data protection that, if implemented, will likely weaken rather than strengthen privacy in the 21st century. Leaving it up to companies and governments to determine the acceptable secondary uses of personal data is a flawed proposition, that will no doubt lead to greater privacy infractions. If the history of privacy has taught us anything, it is that an individual’s loss of control over their personal data leads to greater privacy abuses, not fewer. Inadequate restraints and a paternalistic approach could lead to what privacy advocates fear most — ubiquitous mass surveillance, facilitated by extensive and detailed profiling, sharpened information asymmetries and power imbalances, ultimately leading to various forms of discrimination, old and new.

We recognize that the world is clearly changing. In fact, in the context of online privacy policies in an era of Big Data, Cloud Computing and the Internet of Things, few of us, including the authors of this paper, expect individuals to navigate their way through dense and lengthy privacy notices and policies in order to understand how to protect their privacy. Of course they won’t be able to do that, nor will they be inclined to do so, given the myriad uses of their data, many of which

* Many thanks to Justin Brookman, Director of Consumer Privacy at the Centre for Democracy and Technology, for noting the term “data paternalism,” and Daniel Solove for his reflections on privacy paternalism in “Privacy Self-Management and the Consent Dilemma.” Harvard Law Review. 126 (2013): 1879-2139.
may be collected passively via systems and sensors that are increasingly inhabiting our lives. But we can infer that individuals still have basic expectations of how their personal data will be used, in accordance with the reason(s) why they’re being asked to provide it. The recent opinion of the Article 29 Working Party (WP29) notes that “When we share personal data with others, we usually have an expectation about the purposes for which the data will be used.” This is precisely why placing limits on specifying the purpose, collection and uses of personal information are so important and should ideally be embedded as the default.

The authors of this paper propose a user-centric approach. Over the years, informed and empowered individuals have served as essential checks on the misuses of personal data. In Germany, the concept of informational self-determination was created over 30 years ago by the Constitutional Court who derived it from their Constitution in 1983. It is the perfect way in which to reflect the intent of the OECD FIPPs — that it is the individual — the data subject — who should ultimately determine the fate of his or her personal data.

The paper discusses the recent opinion of the WP29 on purpose limitation, which was intended to protect individuals by restricting how data controllers use personal data, while also providing a degree of flexibility, building on the two elements of purpose specification and compatible use. Purpose Specification is even more critical when individual participation and consent have been diminished. Regardless of consent, individuals will always have basic expectations about how their personal data is to be used, namely that it will be used for the purpose(s) for which they provided it. There is a natural expectation that there will be some basic limitations when you provide your personal data. You don’t hand over your information to the government or to a company to do whatever they wish with it. No – you provide it to fulfill a particular purpose, implicitly or explicitly stated. The term “consistent purpose,” used in the jurisdiction of Ontario, Canada is described and provides an example of how it is applied by a regulator, with parallels drawn to “compatible use.”

The authors fully agree that accountability should be strengthened, but disagree with the proposal to weaken critical FIPPs and diminishing the role of the individual. They argue that (1) diluting consent weakens essential privacy protections; (2) Diminishing limits on specified purposes, collection and uses of personal data minimizes rather than strengthens accountability; (3) privacy requirements are not obstacles to innovation or to realizing societal benefits from Big Data analytics — privacy measures can actually foster innovation and doubly-enabling “win-win” outcomes; (4) greater reliance on law and regulation alone to police “after-the-fact” abuses of personal data is a misguided strategy; and (5) there is little consensus on defining “harms” or ways in which to measure or mitigate privacy harms.

The paper proposes that Privacy by Design principles better reflect current realities and needs by extending the OECD FIPPs, rather than curtailing them. Special mention is made of strong de-identification methods and techniques, which allow innovative and socially beneficial secondary uses of personal data without the need to obtain additional consent, resulting in positive-sum, win-win outcomes. When applied diligently, Privacy by Design extends user controls and enhances accountability, promoting an innovative, design-aware future.
An alarming view is emerging that we believe must be addressed, head on. A proposal was recently published that the Fair Information Practice Principles (FIPPs) contained in the OECD Guidelines — the basis of privacy laws around the world — should be dramatically revised in the era of Big Data, Cloud Computing and the Internet of Things.\footnote{Fred H. Cate, Peter Cullen, and Viktor Mayer-Schönberger, “Data Protection Principles for the 21st Century: Revising the 1980 OECD Guidelines,” December 2013, http://www.oii.ox.ac.uk/publications/Data_Protection_Principles_for_the_21st_Century.pdf.} Co-authors Viktor Mayer-Schönberger, Fred Cate, and Peter Cullen argue that the principles of Purpose Specification, Collection Limitation and Use Limitation should be diminished, in favour of greater emphasis on transparency and accountability by public and private-sector data users/controllers.\footnote{The term “data users” in this paper refers to both data controllers and data processors.} While we are certainly not opposed to greater accountability — far from it — we strongly disagree with the need to diminish critical FIPPs, as recommended by these authors.

Let us be clear: We well recognize that the world is changing. We do not expect individuals to navigate their way through dense and lengthy privacy notices and policies, to somehow understand how to protect their privacy. Of course they won't be able to do that, nor will they be inclined to do so. But individuals have basic expectations that their personal data will only be used for the purposes for which they provided it. “When we share personal data with others, we usually have an expectation about the purposes for which the data will be used.”\footnote{Article 29 Data Protection Working Party, “Opinion 03/2013 on purpose limitation,” April 2, 2013, http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf, p. 4.} This is precisely why limits on specifying the purpose(s) of the collection and use of personal information are so important and should be embedded as the default into information technologies, organizational practices, and networked infrastructures. This user-centric approach contrasts dramatically with the paternalistic approach that we outline below.

Again, let us be clear — there is no expectation that individuals will drop all else and seek out ways in which to consent to the myriad uses of their data, many of which may occur passively via systems and sensors that are increasingly inhabiting our lives. But that does not mean that expectations of limiting the collection and use of personal data to particular purposes should not exist. Quite the contrary — such measures should be embedded as the default, precisely because of the impossible
task of clicking the right opt-out box (if you manage to find it to begin with). Privacy assurance must become embedded, by Design, not left to others to determine.

We chose to reference the concept of paternalism because we believe it reflects the viewpoint expressed in the arguments presented by Mayer-Schönberger et al. Paternalism is defined as: “The attitude or actions of a person, or organization, that protects people and gives them what they need, but does not give them any responsibility or freedom of choice.”\(^4\) We will argue that revising the OECD Guidelines to reduce the principles of Purpose Specification, Collection and Use Limitation, for the “good” of society is misplaced, resting on the fictional notion of an ever-benevolent data user/controller. Taking away an individual’s freedom of choice relating to the specific purposes for which one’s personal data will be collected (purpose specification) and used, does not ultimately benefit the individual — it makes them vulnerable to the judgement exercised by others — corporate and bureaucratic systems that already affect our lives, and over which we have little or no control.


The above paper published by Mayer-Schönberger et al. in December 2013 summarized their five “Priorities for modernizing the OECD Guidelines.” It should be noted that the OECD guidelines were already revised 6 months earlier, in July 2013.

- Reduce the focus on data collection and the attending notice and consent requirements, and focus more on a practical assessment of the benefits and risks associated with data uses.
- Eliminate or substantially reduce the role of the Purpose Specification and Use Limitation principles, which require a specific, articulated purpose for collecting personal data, and restrict data uses to that purpose or related, “not incompatible” purposes.
- Restore the balance between privacy and the free flow of information that was the original goal of the OECD Guidelines, and avoid suppressing innovation with overly restrictive or inflexible data privacy laws.
- Make data users more accountable for the personal data they access, store, and use, and hold them liable when harm to data subjects occurs.
- Adopt a broader definition of the “harms” that inappropriate uses of personal data can cause, and put in place practical frameworks and processes for identifying, balancing, and mitigating those harms.

We disagree with the first three proposals, and have reservations about the last two. While the intent of Mayer-Schönberger et al. may be to shift the burden of privacy protection away from individuals and towards data users/controllers, the effect of their proposals will be to weaken fundamental privacy rights of individuals, while strengthening the power of data users/controllers to decide what personal data to collect and process, whenever and however as they see fit, placing greater burdens on both individuals and regulators to seek effective redress.

Moreover, the proposals refer to the 1980 version of the OECD guidelines. As noted above, a revised set of OECD guidelines was published in July 2013, based on a comprehensive review by The Privacy Experts Group of the OECD Working Party on Information Security and Privacy. Indeed, the OECD members had already identified a number of elements believed to be critical to
improving the effectiveness of privacy protections that included, for example, “embedding privacy by design into privacy management processes.”

In light of Edward Snowden’s revelations of widespread mass surveillance by the state, and with governments also gaining access to large databases in the private sector (as well as the historical record of state abuses), we question the desirability of lowering the standards of privacy and data protection. Quite the opposite — we believe these standards need to be elevated and monitored effectively.

**Diminishing Consent Weakens Essential Privacy Protections and Freedoms**

There are certainly challenges with the prevailing “notice and choice” model to vesting individuals with a right of participation at the time their personal data is collected. Notices are often lengthy and complicated, hard to understand, and inconvenient to access — practical options may be limited. In the emerging Internet of Things, Big Data and Cloud Computing environments, the individual is often unaware of the data collection taking place or may be completely absent from the transaction being processed. In many contexts, providing effective “notice and choice” to individuals about data processing operations may seem like an unnecessary, pointless burden.

Consent — explicit or implied, remains the cornerstone of modern FIPPs and is foundational to modern private sector privacy laws in force around the world. FIPPs are interrelated and intended to be applied holistically. Diminishing consent threatens to diminish an individual's right of participation in the management of one's personal data by others, should he/she wish to do so. In the process, this could also unravel the remaining FIPPs, in their inter-related application. Consent, however obtained or implied, empowers individuals to exercise their privacy rights and freedoms, such as the ability to:

- make consent conditional;
- revoke consent;
- deny consent for new purposes and uses;
- be advised of the existence of personal data record-keeping systems;
- access personal data held by others;
- verify the accuracy and completeness of one's personal data;

---


• obtain explanation(s) of the uses and disclosures of one’s personal data; and
• challenge the compliance of data users/controllers.

Informed and empowered individuals have served as essential checks on the misuses of personal data. In Germany, the concept of informational self-determination was created over 30 years ago by the Constitutional Court who derived it from their Constitution in 1983.\(^9\) It is the perfect way in which to reflect the intent of the OECD FIPPs — that it is the individual the data subject, who should determine the fate of his or her personal data. This captures the central role that the individual is expected to play in determining the uses of his or her personal data. Individuals are intended to feature prominently in considering the acceptable secondary uses of their personal data. Central to this determination is context — context is key to determining what may be considered an appropriate secondary use, and is often lacking without the involvement of the data subject.

Consent-lite regimes will likely fail to be meaningful as there are many instances when the type of collection, especially if directly obtained from the individual, will nonetheless involve consent. Also, in a regime which envisions greater accountability, consent would likely continue to be relied upon by organizations as a precaution against claims of redress.

Removing consent from the equation risks undermining fundamental individual rights, protections and freedoms far beyond “notice and choice” systems. Instead of doing away with consent, we should work on improving transparency and individual control mechanisms — addressing the challenges head-on. Let’s not throw out the consent baby with the data bathwater for the sake of big promises of future benefits. The latter takes the form of a dated zero-sum proposition — that you can only have an increase in one area, at the expense of another. This is an inherently flawed proposition consisting of false dichotomies and unnecessary trade-offs.

**Removing Purpose and Use Limitation is Inversely Related to Accountability**

Eliminating or substantially reducing the basic need to specify purposes and impose justifiable limits on the collection, use and disclosure of personal data gives an unprecedented free hand to data users — public or private, large or small, wherever in the world they may be located, to unilaterally decide why, what, or when personal data should be collected, used and disclosed, with little input from data subjects or oversight authorities.

Lacking sufficient restraints and taking a paternalistic approach could lead to what privacy advocates fear most — ubiquitous mass surveillance, facilitated by more extensive, and detailed profiling, sharpened information asymmetries and power imbalances, ultimately leading to various forms of discrimination.\(^10\) A greater burden would be placed upon both individuals and regulators to prove

---


harms, establish causation, and seek effective redress — the exact opposite of Privacy by Design, which emphasizes prevention and the taking of proactive measures (described below).

If the history of privacy has taught us anything, it is that an individual's loss of control over their personal data leads to greater privacy abuses, not fewer. It is not difficult to imagine how the authors’ proposals, if implemented, could lead to a “collect the entire haystack” mentality, and to overbroad or unspecified and undesirable secondary uses — “fishing expedition” methods of data processing. When making decisions affecting individuals, out-of-date or incomplete data, incorrect inferences, and automated decision-making processes can have profoundly negative consequences. These laissez-faire approaches to data management were not tolerated in the past, so why should we be asked to tolerate them now?

The Purpose Specification principle is even more critical when individual participation and consent have been diminished. Whether or not consent is informed or explicit, individuals will always have basic expectations about how their personal data is to be used, namely, that it will be used for the purpose(s) for which they provided it. There is a natural expectation that there will be some basic limitations when you provide your personal data. You don't hand over your information to the government or a business to do whatever they want with it. No — you provide it to fulfill a particular purpose, implicitly or explicitly stated.

On April 2, 2013, the European Union's Article 29 Data Protection Working Party (WP29) provided an opinion on the principle of purpose limitation. In particular, the WP29 discussed the principle of purpose limitation under the current European Union (EU) Directive 95/46/EC and provided recommendations for the proposed E.U. General Data Protection Regulation.

In the WP29 Opinion, the WP29 stated that purpose limitation protects individuals by restricting how data controllers use personal information, while also providing a degree of flexibility. The WP29 further described purpose limitation as being comprised of two elements: 1) purpose specification; and 2) compatible use. The WP29, explained the relationship between these two elements by referencing Article 6(1)(b) of the E.U. Directive which states that personal information must only be collected for “specified, explicit and legitimate purposes” (purpose specification) and not be “further processed in a way incompatible” with those purposes (compatible use).11

The WP29 also stated the following: “When we share personal data with others, we usually have an expectation about the purposes for which the data will be used. There is a value in honouring these expectations and preserving trust and legal certainty, which is why purpose limitation is such an important safeguard, a cornerstone of data protection. Indeed, the principle of purpose limitation inhibits ‘mission creep’, which could otherwise give rise to the usage of the available personal data beyond the purposes for which they were initially collected. On the other hand, data that have already been gathered may also be genuinely useful for other purposes, not initially specified. Therefore, there is also a value in allowing, within carefully balanced limits, some degree of additional use. The prohibition of ‘incompatibility’ in Article 6(1)(b) does not altogether rule out new, different uses of the data — provided that this takes place within the parameters of compatibility.”12

12 Ibid., p. 4.
The WP29 goes on to state that compatibility needs to be assessed on a case-by-case basis, with the following factors taken into account when assessing compatibility:

- the relationship between the purposes for which the personal data have been collected and the purposes of further processing;
- the context in which the personal data have been collected and the reasonable expectations of the data subjects as to their further use;
- the nature of the personal data and the impact of the further processing on the data subjects;
- the safeguards adopted by the controller to ensure fair processing and to prevent any undue impact on the data subjects.13

Similarly, in the jurisdiction of Ontario, Canada, the Freedom of Information and Protection of Privacy Act (FIPPA) and its municipal equivalent (MFIPPA) limit an institution’s ability to use information in its custody and control. Specifically, section 41(1)(b) of FIPPA and section 31(b) of MFIPPA state that: “An institution shall not use personal information in its custody or under its control except, (b) for the purpose for which it was obtained or compiled or for a consistent purpose.” In determining whether a use is “consistent” with the primary purpose, section 43 of FIPPA and section 33 of MFIPPA provide that a use or disclosure will be considered consistent only if “the individual might reasonably have expected such a use or disclosure.”

In determining whether the individual might reasonably have expected such a use or disclosure, the practice of the Office of Information and Privacy Commissioner of Ontario, Canada (IPC/Ontario) has been to impose a “reasonable person” test. Therefore, the question that must be asked is whether an individual would have reasonably expected the use of their personal information for the identified purposes. Previous investigation reports issued by the IPC/Ontario have found that there must be a rational connection between the purpose of the collection and the purpose of the use, in order to meet the “reasonable person” test.

In applying the “reasonable person” test and determining whether there is a rational connection, IPC/Ontario considers many factors, including the factors listed the WP29 when assessing compatibility.

It is important to note that section 43 of FIPPA and section 33 of MFIPPA define “consistent” purpose in relation to personal information that has been collected directly from the individual. Where information has been collected indirectly, a consistent purpose would be one that is “reasonably compatible” with the purpose for which the personal information had been obtained. Note that IPC/Ontario’s “reasonably compatible” language is virtually identical to the E.U. WP29 “compatible use” language. IPC/Ontario’s practice when assessing “reasonably compatible” purposes is not an “identical purpose” test; rather, IPC/Ontario looks to what the wording and intent of the indirect collection of the information indicates.

It should also be noted that when a consistent purpose cannot be established, Ontario institutions may still use the personal information in their custody or control if the person to whom the information relates has identified that information and consented to its use.14

13 Ibid., p. 3.
14 Freedom of Information and Protection of Privacy Act, RSO 1990, s. 41(1)(a). Please note that section 41(1) of FIPPA and section 31 of MFIPPA specify other purposes for which an institution may use personal information, most of which are beyond the scope of this paper.
As evidenced above, privacy legislation in both the E.U. and Ontario, Canada, place justifiable limits and provide flexibility on a data user’s collection, use and disclosure of personal information. Clearly, these jurisdictions appreciate the dangers of giving data users/controllers the ability to unilaterally decide how personal data should be collected, used and disclosed, with little input from data subjects or oversight authorities.

Privacy Breeds Innovation — Not the Reverse!

Privacy and data protection are at times contrasted with other legitimate societal values and goals, with the suggestion that one area must yield to the other. But is it really necessary to weaken existing privacy measures in the name of pursuing greater efficiencies, innovation and economic growth? No — we believe it is not. The examples of innovative and socially beneficial uses of Big Data analytics that Mayer-Schönberger et al. cite in their paper are already being achieved under current privacy laws. Moreover, as we outline below, Big Data analytics may be pursued using strong de-identification methods and techniques that are fully compatible with privacy. What is often missing in such analytics is a determination of context: What is the appropriate context associated with the data? Context is critical to privacy.

The goal of reconciling privacy rights with the free flow of data was reaffirmed by the OECD in a multi-year review and thorough update of the original 1980 OECD Guidelines; this was completed six months before Mayer-Schönberger et al. published their paper calling for reforms to the OECD Principles. It is noteworthy that all eight of the original OECD Principles were left intact and unchanged (with the exception of using gender neutral language). Thus, the argument that data privacy laws (based upon the OECD Principles) are overly restrictive and suppress innovation does not appear to be shared by the 34 member states who participated in the review process.15

Further, there is a long and growing list of public and private-sector authorities in the United States, the EU, and elsewhere, who unequivocally endorse Privacy by Design as a more robust application of FIPPs, and as a critical means by which to establish sufficient, necessary trust in the evolving information economy.16 Privacy by Design was unanimously endorsed as an international framework for privacy by the International Congress of Privacy Commissioners and Data Protection Authorities in 2010. Privacy by Design promotes prevention and innovation, resulting in doubly-enabling, positive-sum outcomes, and has now been translated into 35 languages.17

16 These include, inter alia, the U.S. White House, Federal Trade Commission, Department of Homeland Security, Government Accountability Office, European Commission, European Parliament and the Article 29 Working Party, among other public bodies around the world who have passed new privacy laws based upon the FIPPs. In addition, international privacy and data protection authorities unanimously endorsed Privacy by Design as an international standard for privacy.
More Accountability, Definitely! Relying on After-the-Fact Redress, Why?

We are concerned that Mayer-Schönberger et al. propose holding data users liable for the actual harms that occur to individuals. Everyone agrees that greater accountability for the uses of personal data is critical.\footnote{Indeed, important work has been carried out in this area in recent years by the OECD, the E.U. Commission, the FTC in the United States, and many other public and private sector industry associations, standards-setting bodies and advocacy groups.} However, this proposal shifts the burden of proof to demonstrate the existence of harm to individuals, with regulators officiating such cases to document the harms, to prove causality, and then seek redress. Proving the causality of harms is notoriously difficult to do, and will likely become even more so in the current era of complex, interconnected global information systems and networks that are increasingly opaque to both individuals and oversight authorities.

Even today, harms arising from cases of identity theft due to a security breach are difficult to prove. Similarly, establishing links between poor organizational data-handling practices and the negative effects of individuals being erroneously placed on a watchlist or other similar blacklist, losing an employment opportunity, paying a higher insurance premium, being denied health coverage, or suffering a damaged reputation or the inability to travel, can be a Kafkaesque experience.

While superficially appealing in theory, in practice, harms tests are far too narrow a basis for effectively protecting privacy in this day and age.\footnote{See Ryan Calo, "The Boundaries of Privacy Harm," Indiana Law Journal 86, no. 3 (2011): 1131–1162, http://ilj.law.indiana.edu/articles/86/86_3_Calo.pdf. See also the remarks of Marc Rotenberg in Federal Trade Commission Roundtable Series 1 on: Exploring Privacy, December 7, 2009, http://www.ftc.gov/sites/default/files/documents/public_events/exploring-privacy-roundtable-series/privacyroundtable_dec2009_transcript.pdf, p. 301; the remarks of Leslie Harris, ibid., pp. 36–38; and the remarks of Susan Grant, ibid., pp. 38–39.} As the name implies, harms tests are fundamentally reactive, allowing harms to arise rather than proactively preventing the harm, right from the outset. The effect of Mayer-Schönberger et al.’s proposal will be to retard the development and application of real, effective preventative remedies. In the meantime, a mountain of unnecessary harms will have occurred, responsibility for which will most likely go undetected and unchallenged. Have we learned nothing from the past? A flexible, robust set of fair information practices, ideally embedded into design, remains the best bulwark against future harms (material or immaterial). Moreover, regulators’ resources are already stretched to the limit, and it is highly unlikely that additional staffing will be provided to absorb the additional burdens imposed by such a proposal. The opposite is happening — resources are shrinking, not expanding.

No Consensus on “Harm” or Agreed-upon Ways to Measure or Mitigate “Harms”

Even if a harms-based approach to privacy was feasible, we are a long way from achieving meaningful national, let alone international, consensus on defining “harms” (nor broadening the scope). We are far from… “put[ting] in place practical frameworks and processes for identifying, balancing,
and mitigating those harms. And who would do this? U.S. courts have been reluctant to step in on behalf of affected individuals.

Absent clearly defined and agreed upon standards for privacy-related “harms,” any proposal to liberalize the market for collecting, using and disclosing personal data should be viewed with skepticism. Adopting a broader definition of harm may be helpful. Daniel Solove’s excellent taxonomy of privacy was a seminal work that broke new ground in how to conceptualize privacy and ways to mitigate negative impacts, at least from a legal perspective. There should also be greater emphasis on applying risk-based methods and comprehensive privacy impact assessments (PIAs), but here too, standards remain in their infancy, with enormous variation in the approaches taken.

As we noted above, individuals would be significantly disadvantaged by the lack of notice and consent, and the minimization of their ability to participate in the process. Any significant loss of individual autonomy in relation to one’s personal data should be viewed as harmful.

---


Privacy by Design Innovates by Extending User Control

We agree that the prevailing “Notice and Choice” model has many flaws and needs to be strengthened towards a more robust user-centric “Transparency and Control” model. The 7 Foundational Principles of Privacy by Design is such a model (see Appendix). The User-Centric principle encourages innovation in this area, for example, by furthering the “SmartData” concept, which automatically restricts secondary uses within user-centric devices. Trusted online agents and third parties would minimize the creation and processing of personal data automatically, acting as intermediaries and enforcers of individual privacy preferences. Such systems, based on Privacy by Design, promise to extend the ability of individuals to exercise meaningful control over their personal data.

We readily acknowledge that there is much room for innovation to address the needs of an evolving world where individuals are acting less and less as direct parties to online transactions; as such, they have less opportunity to exercise meaningful participation in the lifecycle of their personal data. Considerable work on user-centric, privacy-enhancing and transparency-enhancing technologies is being undertaken by leading E.U. and U.S. researchers, and is deserving of greater attention and support.

23 “User” here refers to the data subject.
25 Cavoukian, “Privacy in the Clouds.”
In addition, the Personal Data Ecosystem (PDE) is an emerging trend supported by a number of companies and organizations that have developed tools and technologies to enable the individual to have much greater management and control over his/her personal information than is currently possible today.27

*Privacy by Design* places the onus upon data users/controllers to anticipate and acknowledge the individual’s privacy interests, wherever possible. An essential *Privacy by Design* principle is Privacy as the Default which specifies that data users should engineer information technologies, organizational processes and networked systems with the most privacy-protective default settings. This is essentially an opt-in model involving the individual’s positive consent for additional secondary uses of their data.

**Privacy by Design Enhances Accountability — Across the Board**

As stated above, we wholeheartedly agree that accountability should be strengthened. There are many ways to achieve this using a *Privacy by Design* framework.29 No less than five of the *Privacy by Design* Principles relate to improving accountability on the part of data users:

- **User-Centric** — Accountable directly to the individual data subject;
- **Keep it Open** — Accountable to the public and to regulators;
- **Proactive** — Accountable within organizations, internal and external;
- **Embedded** — Accountable to business partners and auditors by adopting systematic, privacy protective methods, embedded in design, that may be independently verified; and
- **Positive-Sum** — Accountable to the public, industry and regulators by openly publishing advanced methods and outcomes of achieving privacy (along with other functionalities) for others to learn, adopt, and become best practices.

*Privacy by Design* enhances accountability to individuals, oversight authorities, business partners, shareholders, internal teams, and the public at large. For all the emphasis Mayer-Schönberger *et al.* place on evolving FIPPs and improving accountability, it is a mystery why they made no reference to *Privacy by Design*’s advances in this area. The association of *Privacy by Design* with accountability

---


Towards a Design-Aware Future

We agree that Cloud Computing, the Internet of Things, and Big Data analytics are all trends that may yield remarkable new correlations, insights, and benefits for society at large. While we have no intention of standing in the way of progress, it is essential that privacy practitioners participate in these efforts to shape trends in a way that is truly constructive, enabling both privacy and Big Data analytics to develop, in tandem.31

There is a growing understanding that innovation and competitiveness must be approached from a “design-thinking” perspective — namely, viewing the world to overcome constraints in a way that is holistic, interdisciplinary, integrative, creative and innovative. Privacy must also be approached from the same design-thinking perspective. Privacy and data protection should be incorporated into networked data systems and technologies by default, and become integral to organizational priorities, project objectives, design processes, and planning operations. Ideally, privacy and data protection should be embedded into every standard, protocol, and data practice that touches our lives. This will require skilled privacy engineers, computer scientists, software designers and common methodologies that are now being developed, hopefully to usher in an era of Big Privacy.

We must be careful not to naively trust data users, or unnecessarily expose individuals to new harms, unintended consequences, power imbalances and data paternalism. A “trust me” model will simply not suffice. Trust but verify — embed privacy as the default, thereby growing trust and enabling confirmation of trusted practices.

The Enormous Value of De-identification

Many Big Data applications may be achieved using de-identified data in place of identifiable personal information. De-identified data is personal information from which identifying characteristics have been removed or obscured so that it is not reasonably likely that the data could identify an individual. De-identified data can be used to build models to detect anomalous individuals or to classify individuals into categories (for marketing purposes). De-identification significantly reduces the risk that personal information will be used or disclosed for unauthorized or malicious purposes. In many jurisdictions, it also allows data to be used for secondary purposes, without the need to go back to the data subject for consent.

Dr. Khaled El Emam, Associate Professor at the University of Ottawa and Canada Research Chair in Electronic Health Information, has developed a tool that de-identifies personal information in a manner that simultaneously minimizes both the risk of re-identification and the degree

---

of distortion to the original database. The application of this tool to any database of personal information provides the highest degree of privacy protection, while ensuring a level of data quality that is appropriate for the secondary use. This privacy-enhancing technology provides an excellent example of what may be achieved using a doubly-enabling, positive-sum approach which maximizes both goals — in this case, individual privacy and data quality.

One excellent example is the area of health research: de-identification is particularly valuable in the context of personal health information. Health information is highly sensitive and may include some of the most intimate details associated with one’s life, such as those related to one’s physical or mental health. Personal health information requires the strongest privacy and security protections to prevent its unauthorized collection, use and disclosure. However, under appropriate circumstances, it is critical to provide access to this information for vital secondary purposes that are strongly in the public interest. For example, health information is essential for public health purposes and health-related research. It is also used for purposes such as planning, delivering, evaluating and monitoring health programs and services, and improving the quality of care. The availability of information for such purposes results in enormous benefits for individuals and society at large by improving health-care programs and services and by improving the effectiveness of the health-care system. Health research can provide critical information about disease trends, risk factors, outcomes of treatment, and patterns of care — it has led to significant discoveries, including the development of new treatments and therapies.

Practical de-identification is risk-based. The amount of transformation to a database to protect it is contingent on the context. Factors such as the security and privacy controls that a data user has in place, any restrictions in contracts, and information sensitivity are accounted for in this risk assessment. There is evidence that when best practices of de-identification are followed, the risk of re-identification is rendered extremely minimal. If data users routinely de-identified personal information, there would also be far fewer data breaches and cases of identity theft.

A non-trivial percentage of data breaches arise from “inside” jobs — by rogue employees who have easy access to identifiable data, or accidentally by employees who do not follow good data management practices. Such breaches could be reduced dramatically by default if far less personal data were retained in identifiable form — instead, being routinely retained with an appropriate amount of de-identification applied. The routine de-identification of information would also help to prevent privacy breaches in cases where the media storage devices were lost, stolen or accessed by unauthorized third parties.

If it is necessary to re-identify individuals, then a pseudonym may be used to re-identify specific individuals that require further attention. This is what typically happens in public health studies when done in a privacy-preserving manner: the analysis is done on de-identified or pseudonymized data with specific individuals only re-identified for health-related purposes. It is important to note, however, that from a European and German law perspective, such de-identified data would remain as personal information because the individual may later be re-identified under certain conditions. As such, it would fall within the scope of privacy laws.

De-identification, pseudonymisation and anonymisation are vitally important tools to protect privacy in the context of free access to information. The European Data Protection Commissioners have developed criteria and practical guidelines on open data and public sector information re-use, as has the Office of the Information and Privacy Commissioner of Ontario, Canada.

---


In this paper, we have set out our objections to the proposal that Purpose Specification, Collection and Use Limitation be abridged in order to allow for Big Data and other technological innovations — we view this as fundamentally zero-sum thinking. We wish to direct this message to all those who would argue that privacy principles prevent much-needed and altruistic uses of data, in order to advance societal interests. Not only could the indiscriminate collection of personally identifiable data cause irreparable harm to individuals, but such practices may impede much sought-after progress in the sciences, health sector, and education. For example, in the case of Big Data, one may argue for the need to “gather the haystack” in order to “find the needle,” when in reality, it could be much easier to find the needle without the haystack. The default cannot be “collect all the data” in personally identifiable form. Privacy should be the default setting. But within that context, great strides may be made in data science and Big Data analytics. This is not an either/or proposition — abandon zero-sum thinking.

Our view is that the OECD principles should remain inherently intact and may be further enhanced through the application of Privacy by Design, which adds new elements to traditional FIPPs, such as proactively embedding privacy into information technologies, business practices, and network infrastructures. By doing so, individuals are not placed in the position of having to be concerned about safeguarding their personal information — they can be confident that privacy is assured, right from the outset. As noted above, many technological applications that Mayer-Schönberger et al. cite could have been achieved without the use of any personally identifiable information. One of the essential tools that enables both Big Data and Big Privacy is the use of strong de-identification techniques (and other techniques that will follow), which minimize the risks and prevent the privacy harms from arising.

We encourage everyone to join us in shaping a future where privacy and innovations such as Big Data and the Internet of Things can intermingle, live and breathe together. Let us abandon zero-sum thinking in favour of doubly-enabling positive-sum systems — we can and must have both!

---

The 7 Foundational Principles of *Privacy by Design*

1. **Proactive** not Reactive; **Preventative** not Remedial

The *Privacy by Design (PbD)* approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events *before* they happen. *PbD* does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred — it aims to *prevent* them from occurring. In short, *Privacy by Design* comes before-the-fact, not after.

2. Privacy as the **Default Setting**

We can all be certain of one thing — the default rules! *Privacy by Design* seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy — it is built into the system, *by default*.

3. Privacy **Embedded** into Design

*Privacy by Design* is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality.

4. Full Functionality — **Positive-Sum**, not Zero-Sum

*Privacy by Design* seeks to accommodate all legitimate interests and objectives in a positive-sum “win-win” manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. *Privacy by Design* avoids the pretense of false dichotomies, such as privacy *vs.* security, demonstrating that it is possible to have both.
5. End-to-End Security — Full Lifecycle Protection

Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire lifecycle of the data involved — strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, secure lifecycle management of information, end-to-end.

6. Visibility and Transparency — Keep it Open

Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify.

7. Respect for User Privacy — Keep it User-Centric

Above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric.
Bibliography


“Privacy and Identity Management for Community Services (PICOS)”, www.picos-project.eu

“Privacy and Identity Management for Europe (PRIME)”, www.prime-project.eu

“Trustworthy Clouds Privacy and Resilience for Internet-Scale Critical Infrastructure (TClouds)”, www.tclouds-project.eu


Ann Cavoukian, and Drummond Reed, “Big Privacy: Bridging Big Data and the Personal Data Ecosystem through Privacy by Design” http://www.ipc.on.ca/english/Resources/Discussion-Papers/Discussion-Papers-Summary/?id=1352

Ann Cavoukian, and Justin Weiss, “Privacy by Design and User Interfaces: Emerging Design Criteria — Keep It User-Centric” http://www.ipc.on.ca/English/Resources/Discussion-Papers/Discussion-Papers-Summary/?id=1201


Merriam-Webster, “Paternalism” http://www.merriam-webster.com/dictionary/paternalism


Ann Cavoukian, Ph.D.
Information and Privacy Commissioner,
Ontario, Canada

2 Bloor Street East, Suite 1400
Toronto, Ontario
Canada M4W 1A8

Web site: www.ipc.on.ca
Privacy by Design: www.privacybydesign.ca
Telephone: 416-326-3333
Fax: 416-325-9195

March 2014