Foreword: Privacy Is Not Quite Like the Weather

Privacy, like the weather, is something everyone talks about.[1] But unlike the weather, there is much that should, and can, be done about it. This welcome volume documents and explains an important tool for doing that. It should be in the library of any professional concerned with collecting, processing, using or determining the fate of personal data (whether as policy-setter, administrator or researcher). This state-of-the-art book describes the most comprehensive tool yet available for policy-makers to evaluate new personal data information technologies before they are introduced.[2]

Privacy impact assessment aims to contribute to organisational practice, as well as culture. It recognises that machine-processed data on persons requires special protections, particularly when new tools are involved. It anticipates problems, seeking to prevent, rather than to put out fires. The PIA model is based on avoiding future problems by learning from the past and imagining how new technologies might bring new problems – including that intriguing class of the "unknown unknowns".

PIA is very much a work in development and offers a general model whose content needs to be adjusted depending on the specifics of the case. One size does not fit all. The variation these chapters consider precludes a standard form, at least with respect to substance and the range and degree of attention to potential problems.

Details of the assessment and expectations of what is appropriate will vary depending on the institution/organisation and goals – private vs. government (and within government national security and crime control as against education or welfare); the role played; location as in geographical co-ordinates or in public or private places); whether data is immediately accessible or requires technical enhancements or otherwise being pried out or constructed; the kind of data – sensitive vs. non-sensitive and intimate vs. non-intimate personal information; the tool and fullness of the form of data it offers (audio and video documenting behaviour vs. merely noting location); identification – unique, masked, fully anonymous or group identification;

[1] These observations draw from various articles at www.garymarx.net and Marx, forthcoming.

[2] Of course, as several of the articles note, a perennial problem and trade-off is intervening too early or too late. Runaway trains can't very well be called back, even as those who build fast trains need the freedom to experiment.

and the fate of the data – is it shared with the subjects, is it sealed, destroyed or available to the public; and the costs of trying to prevent a risk relative to its seriousness and the likelihood of its occurring. It will also vary at different stages of data collection, analysis and use, and for local historical, cultural and social factors.

What is being assessed? Most of the chapters in the volume use privacy broadly to refer to information pertaining to an individual, particularly as it is machine-processed. It begins when the borders of the person are crossed to either take information from or impose it upon a person. Privacy is a general term and there are endless arguments about what it applies to and if it is the best term to capture contemporary concerns.

Most of the authors in this book are implicitly using the form of information privacy identified by Westin[3] – this emphasises control by the subject. This implies an individual right and the actors' ability to make choices. The assumption is that individuals will be well served by a policy when they decide for themselves what personal information to release. What matters is choice and treating the data in accord with Fair Information Practices. That is admirable, but it leaves untouched other important issues.

Applying the conventional principles for the machine processing of information just to information privacy will seem too narrow for many observers. Other issues of great salience for citizens and society are slighted such as the implications for social stratification; for fairness (when the choices are specious or not equally available); for human rights and for silently creating creeping precedents that might lead to unwanted results. Other forms of privacy may also be ignored.

Noting this limitation, Paul De Hert (Chapter 2) considers the need for assessments concerned with human rights more broadly that privacy may (or may not directly) connect with. Raab and Wright (Chapter 17) discuss extending assessments to take more explicit account of various surveillance activities that may touch privacy but are not synonymous with it in its narrow sense.

Whether privacy is the best term to apply to current personal data and new surveillance issues is subject to debate. In an informative exchange in *Surveillance & Society*,[4] Colin Bennett acknowledges the limitations of the concept but makes a strong case for using privacy as a catch-all term for a variety of relevant information issues beyond itself. In popular culture and for interest groups, the term is becoming inclusive of an array of data issues that may connect to privacy, but go far beyond it.

This discussion leads us to ask: what is the PIA tool intended to prevent, or alternatively, what goals to drive forward? What does (and should) the assessment assess

[3] Westin, Alan, *Privacy and Freedom*, Atheneum, New York, 1967.

[4] The debate on the value of privacy in surveillance studies was initiated by Colin Bennett's essay "In Defence of Privacy", *Surveillance & Society*, Vol. 8, No. 4, 2011, pp. 485–496. Respondents, in the same issue, were Priscilla M. Regan ("A Response to Bennett's 'In Defence of Privacy'", pp. 497–499), John Gilliom ("A Response to Bennett's 'In Defence of Privacy'", pp. 500–504), danah boyd ("Dear Voyeur, Meet Flâneur... Sincerely, Social Media", pp. 505–507) and Felix Stalder ("Autonomy beyond Privacy? A Rejoinder to Colin Bennett", pp. 508–512). The debate can be downloaded as a single file: http://www.surveillance-and-society.org/ojs/index.php/journal/article/downloadSuppFile/privacy_defence/privacy_debate

and why? This is forejudged to a degree by using the term privacy. But that choice can be problematic since the latter is such a general concept and can refer to such varied phenomena. Some clarification of terms private and public and surveillance may be helpful.

Untangling Terms

> *If this [dissemination of FBI criminal history records] is done properly, it's not a breach of privacy.*
> *Clarence Kelley, FBI Director*[5]

Privacy is related to a broader family of terms such as publicity and surveillance and anonymity and secrecy. If PIA is to be an effective tool, there is need for a broad and systematic view of the setting and for conceptual differentiation in terminology.

How do surveillance and privacy relate? Surveillance is often wrongly seen to be the opposite of privacy. Kelvin emphasised this role of privacy as a nullification mechanism for surveillance.[6] But at the most basic level, surveillance is simply a way of discovering and noting data that may be converted to information. This obviously can involve invasions of privacy as with the employee in a lab testing for AIDS who sold information on positive results to a mortuary.

Yet surveillance can also be the means of protecting privacy. Consider biometric identification and audit trails required to use some databases, or defensive measures such as a home security video camera. Privacy for whom and surveillance of whom and by whom and for what reasons need to be specified in any assessment.

Depending on how it is used, active surveillance can affect the presence of privacy and/or publicity. As nouns, the latter can be seen as polar ends of a continuum involving rules about withholding and disclosing, and seeking or not seeking, information. Depending on the context, social roles and culture, individuals or groups may be required, find it optional, or be prohibited from engaging in these activities, whether as subjects or agents of surveillance and communication.

The right to privacy can be matched by a right to publicity. There might even be a need for *Publicity Impact Assessments* to be sure that personal information is collected and when appropriate given access to a wider public.

Such assessments would be sure that surveillance and/or communication of results are *mandated* rather than prohibited! One form involves a right to know as with freedom of information rules. Another form can be seen in the right to be acknowledged and noted, implied in the idea of citizenship, for example, in being entitled to have a driver's licence, register to vote or obtain a passport or a national

[5] U.S. News and World Report, 15 October 1973, p. 59.

[6] Kelvin, P., "A Social-Psychological Examination of Privacy", *British Journal of Social and Clinical Psychology*, Vol. 12, 1973, pp. 248–261.

identity card.[7] In some ways, this is the reverse of an expectation not to be defamed or lied about.

When the rules specify that a surveillance agent is not to ask certain questions of (or about) a person and the subject has discretion about what to reveal, we can speak of *privacy norms*. When the rules specify that information must be revealed by the subject or sought by the agent, we can speak of *publicity norms* (or better perhaps *disclosure norms*). The subject has an obligation to reveal and/or the agent has an obligation to discover. With publicity norms, there is no right to personal privacy that tells the agent not to seek information, nor that gives the subject discretion regarding revelation. Rather there is the reverse – the subject has an obligation to reveal and/or the agent to discover. This also suggests a way of broadening assessments regarding personal data. Here the goal is visibility rather than data protection.[8] A source of confusion in discussions of both privacy and publicity involves the failure to differentiate these as adjectives from nouns.

Private and Public as Adjectives

Information as a normative phenomenon involving moral expectations (whether for protection or revelation and whether based on law, policy or custom) can be differentiated from the actual empirical status of the information as known or unknown. For this, we need the related terms private and public – adjectives that can tell us about the status of information. Whether information is known or unknown has an objective quality and can be relatively easily measured. For example, in face-to-face encounters, the gender and face of a stranger is generally known, regardless of whether this is in the street, an office or a home.[9] The information is "public", as in readily accessible.[10] In contrast, their political or religious beliefs are generally invisible and unknown.[11]

Of course, normative expectations of privacy and publicity do not always correspond to how the adjectives *public* and *private* are applied to empirical facts. Thus,

[7] The Spanish Data Protection Agency in its justifying Spain's new mandatory national identity card claims that the card goes along with the citizen's right to a national identity (Ouzeil, Pablo, 2010)

[8] Marx, Gary T., "Turtles, Firewalls, Scarlet Letters and Vacuum Cleaners: Rules About Personal Information", in W. Aspray and P. Doty (eds.), *Making Privacy*, Scarecrow Press, Lanham, MD, 2011 [forthcoming].

[9] There may, however, be rules about the subsequent recording, communication and use of such information.

[10] Identification may be controlled through anti-mask laws or conversely through requiring veils for females or the display of religious or other symbols (tattoos, brands, badges) or clothing indicating status.

[11] Revelation, of course, may be mandated by rules requiring the wearing of symbols indicating these. This paragraph also assumes that in most cases people "are" what they appear to be, i.e., no cross-dressing.

the cell phone conversations of politicians and celebrities that have privacy protections may become public. Information subjected to publicity requirements such as government and corporate reports and disclosure statements may be withheld, destroyed or falsified. Information not entitled to privacy protections, such as child or spouse abuse, may be unknown because of the inaccessibility of the home to broader visibility. The distinction here calls for empirical analysis of the variation in the fit between the rules about information and what actually happens to it.

Privacy and publicity can be thought of in literal and metaphorical spatial terms involving invisibility-visibility and inaccessibility-accessibility. The privacy offered by a closed door and walls and an encrypted e-mail message share information restriction, even as they differ in many other ways. Internet forums are not geographically localised, but in their accessibility can be usefully thought of as public places, not unlike the traditional public square where exchanges with others are possible.

There would be more agreement, or at least greater clarity, if assessments of privacy were clearer about whether they are talking about respect for the rules protecting privacy or the empirical status of information as known or not known. When the laws are followed, former FBI director Clarence Kelley (in the quote that begins this section) can correctly claim that they haven't been breached with respect to privacy. But he could not claim that, as an empirical matter, privacy is not altered when such records are created and circulated.[12]

Types of Privacy

Privacy is a multi-dimensional concept with fluid and often ill-defined, contested and negotiated contours, depending on the context and culture. PIAs should be clear about what privacy means for the context with which they are concerned (and often more than one meaning will apply).

Within informational privacy with which the chapters here are largely concerned, we find the conditions of anonymity and pseudo-anonymity, often referred to as being necessary for another type of privacy involving seclusion and being left alone. Personal borders are obviously more difficult to cross if an individual cannot be

[12] There can, of course, be verbal prestidigitation, not to mention bad faith, in simply defining away invasions of privacy as non-existent because the law or rules are followed. The deeper issue is what degrees of control does the individual have over personal and private information and are the lines appropriately drawn given a society's values and broader transcendent values of human dignity and life.

There are also other sources of confusion such as the legal definition of geographical places and information as public or private, custom and manners (e.g., averting the eyes) and roles which offer varying degrees of access to information. See Marx, Gary T., "Identity and Anonymity: Some Conceptual Distinctions and Issues for Research", in J. Caplan and J. Torpey, *Documenting Individual Identity*, Princeton University Press, 2001.

reached via name or location. The conditions around revelation or protection of various aspects of identity are central to the topic.

Informational privacy encompasses physical privacy. The latter can refer to insulation resulting from natural conditions such as walls, darkness, distance, skin, clothes and facial expression. These can block or limit outputs and inputs. Bodily privacy is one form of this. This is seen in crossing the borders of the body to implant something such as a chip or birth control device or to take something from it such as tissue, fluid or a bullet.[13]

A related, taken-for-granted form is aesthetic privacy[14] which refers to the separation, usually by a physical barrier of bedroom or bathroom, of activities involving one's "private parts" and unguarded moments. Alderman and Kennedy discuss a number of such cases in which the shock of discovering a violation surfaces norms of which we are hardly aware because they are so rarely violated.[15] Clothes and manners also sustain this. The concern over full body airport scans is also illustrative.

Informational privacy can be further descriptively considered as it ties to institutional setting (e.g., financial, educational, health, welfare, employment, criminal justice, national security, voting, census); places and times; the kind of data involved such as about religion or health, apart from the setting; participant roles (communications privacy as involving two-party, one-party or no-party consent); and aspects of the technology such as wire or wireless, phone, computer, radio or TV. PIAs need to consider setting, data type and means –factors that are central to legislation and regulation and rich in anomalies.[16]

In emphasising informational privacy, several other commonly considered forms such as decisional[17] or proprietary[18] privacy are slighted. These primarily involve application or use, rather than information discovery.

Defining cases in the US such as *Griswold* v. *Connecticut* 381 U.S. 479 (1965) and *Roe* v. *Wade*, 410 U.S. 11 (1973) involve decisional privacy with respect to personal and intimate matters such as family planning, birth control, same sex marriages or physician-assisted suicide. Proprietary privacy – use of a person's

[13] The physical border perspective has limits too, thus taking/giving a urine, breath sample or photo involves using things that have already left the body and are different and beyond the literal physical protective border. The situation is the same for garbage. The borders in such cases are cultural – note the tacit assumption that one's garbage isn't to be examined – at least in a personally identifiable way.

[14] Rule et al., 1980.

[15] Alderman, Ellen, and Caroline Kennedy, *The Right to Privacy*, Alfred A. Knopf, New York, 1996.

[16] Thus, in the US, the Federal Communications Commission has jurisdiction over content delivered over a wire but not that by satellite. In countries such as Germany and France, privacy rights are defined in reference to broad constitutional principles such as the dignity of the person, while in the US, the particular technology or institution plays a much larger defining role.

[17] DeCew, Judith, *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology*, Cornell University Press, Ithaca, NY, 1997.

[18] Allen, 2007.

information without consent for commercial and other purposes also involves control and liberty questions and the extension of market principles to symbolic material that is often immaterial (at least physically).

While distinct, informational privacy shares with the other forms inclusion in the broader category of privacy as control over access to the person or at least the person's data, these may be connected. Thus, if individuals can control their personal information – whether not having to reveal their purchase of birth control pills (when this was illegal) or keeping paparazzi from taking pictures, then they need not worry about that information being used.

In addition to conceptual elaboration and reflection on the broader consequences of data collection, privacy assessment needs can be analysed in more detail when protective activities and problem avoidance are connected to a series of sequential stages that can be observed in the process (and processes) of data generation and use. Various types of privacy problem do not occur randomly, but tend to cluster at particular social locations.

Data Stages

Privacy protection is not like a vaccination that occurs once and is over. Rather it is part of an enduring process involving a series of separate actions.

Table 1 lists seven kinds of activity called *surveillance strips* that follow each other in logical order. The strips are temporally, conceptually, empirically and often spatially distinct.

Over time, the distinct action fragments of these stages combine into stories about personal data and illustrate the emergent character of surveillance and privacy as multi-faceted abstractions made up of many smaller actions. These are not unlike the frames in comic books (although not intended to be entertaining and the patterns are more like the fluid, jumpy sequences of cyberspace explorations than the rigid frame ordering of the comic book).

When viewed sequentially and in their totality, these elements constitute *surveillance occasions*.[19] A surveillance occasion begins when an agent is charged with

Table 1 Seven surveillance strips

(1) tool selection
(2) subject selection
(3) data collection
(4) data processing/analysis [raw data] numerical/narrative
(5) data interpretation
(6) uses/action
(7) data fate (other)

[19] Goffman discusses strips (1964) and occasions (1974) in referring to face-to-face interaction. See Goffman, Erving, *Behavior in Public Places: Notes on the Social Organization of Gatherings*, The Free Press, 1963, and *Frame Analysis: An Essay on the Organization of Experience*, Harper

the task of gathering information. Following that, the seven phases in Table 1 can be considered.[20] Studying the behavioural sequences of tool selection, subject selection, data collection, data processing, interpretation, resulting action (or inaction) and fate of the data offers a way to order the basic behaviours occurring within the family of direct surveillance actions.[21] The stages are the direct pressure points where most problems will be found.

Sometimes these occur almost simultaneously as when a motion sensor is triggered, a message is sent to a central computer, an alarm sounds and a door is locked or a retinal pattern is matched to a given identity and a computer unlocks or when a video camera does not save what unproblematic passes before it. But for a goodly proportion of applications, as with drug testing or data mining, these consist of different activities and stages and involve a division of labour with agent roles played by various actors.

In a given story, the stages may develop in a serial fashion as one stage logically leads to the next (e.g., from data collection to analysis) or it may stop early on (a tool and subject are identified but no data is collected, or the data is not analysed or applied). Looking across many cases suggests a variety of ideal-type career patterns (with different stopping and turning points). However, once data has been gathered, questions regarding the data's fate (the last item in the chain) can always be asked.

Most of the privacy and related problems with which PIAs are concerned occur (when present) at one of the stages in Table 1. The kind of problem may differ by stage – thus violations of consent are likely at data collection, of fairness and validity at processing and interpretation, of discrimination at use and of confidentiality at data fate.[22]

---

and Row, London, 1974. However, as used here, they refer to bundles of discrete activity from the point of view of the observer and most do not involve face-to-face interaction of agents and subjects.

[20] Decisions about *who* is responsible for doing the surveillance and the design of the technology could be treated as the initial strips as well. However, attention here is on the next stage directly associated with doing the surveillance.

[21] This is said mindful of the fact that it is always possible to make ever greater differentiations within the categories identified and to push the causal chain back farther. For example, with respect to the data collection phase, contrasts can be made based on the tool, the sense involved, the kind of activity or the goal. The Table 1 conceptualisation captures the major natural breaks in activity once a problem in need of personal information has been defined and an agent designated.

[22] It would be useful to have a checklist of problems that can occur and (when possible) of ways of avoiding them, or ameliorating them when they can't be prevented. The list would include various kinds of physical, psychological and social harm and unfairness in application and use; minimising invalid or unreliable results; not crossing a personal boundary without notice or permission (whether involving coercion or deception or a body, relational, spatial or symbolic border). Other problems to be avoided involve violating trust and assumptions that are made about how personal information will be treated (e.g., no secret recordings, respect for confidentiality, promises of anonymity or for the compartmentalisation of kinds of data). Lists are one thing and can be a bit like waving the flag. Who, after all, would favour invalid results or violating trust? There is a strong need for research on the frequency and social locations of such problems. When are they common, patterned and systemic as against being infrequent, random and idiosyncratic?

Knowledge and policy are better served when these elements are differentiated. The stages in Table 1 can direct research on the correlates and location of particular kinds of problems. Awareness of the stages of the process can help in assessing the seriousness and likelihood that a risk will occur and the costs of prevention (whether by not using, regulating or amelioration after the fact).

The likelihood of prevention is also greatly affected by the stage. Just saying "no" to a data collection request (if honoured) is the ultimate prevention. But as the process moves from collection to the final fate, controls become more challenging. In the initial stages, the relevant actors and locations for accountability are known – but over time, if the information spreads out in wider circles, as it often does, control weakens.

Slices, Not Loaves

If the wise suggestions this specialised volume recommends were implemented, there would be far fewer problems associated with the collection and processing of personal data. However, the authors are hardly naïve reformers promising salvation if only their preferred solutions are followed. Limits are identified, as are ways of working within or around many of them.

Sceptical pundits removed from any responsibility for action can, of course, snipe from the sidelines about PIAs and their limits. PIAs are generally not mandated, requiring voluntary introspection and self-restraint on the part of goal-focused (often bottom line) organisations.[23] Businesses are not democracies and government's national security and crime functions require levels of secrecy. When a PIA is carried out, results may not become public. Will a PIA's requirements be implemented? Or will PIAs serve as window dressing disingenuously prohibiting, while hiding behaviour that would be unacceptable if made public? Will they become another ritualised hurdle to jump over (or under) for busy practitioners with more important goals?

In an effort to learn and to legitimate, the ideal PIA involves relevant "stakeholders". This democratic impulse is admirable, but who decides who is a legitimate stakeholder? –e.g., do those arrested, but not charged or found guilty, have a seat at the table when decisions are made about preserving DNA? Do free speech as well as privacy advocates serve on a telecommunications committee charged with assessing a new technology?[24]

PIA faces the challenge of preventing a particular kind of future which involves new elements. It goes beyond routine audits of compliance with established rules

[23] There are a few exceptions as several chapters here note such as in Europe for RFID and in the US for e-government.

[24] The failure to not include types of consumer in the decision to roll out caller-ID created problems for US telephone companies in the 1980s. Those with unlisted numbers and shelters for abused women lost control over their phone numbers by technological fiat and there was much public outcry which led to revised policies.

and policies. Since the future hasn't yet happened, its assessment is forever vulnerable to challenges and doubts.

Given still other challenges – from political pressures to lack of resources, it is noteworthy that the stellar policy analysts in this book have not given up. They are thoughtful realists – dealing humbly, yet hopefully, with terribly complicated contemporary questions. In situations drenched in trade-offs, legitimate conflicts of interest and uncertainty, the lack of a full loaf should not be bemoaned, rather one should be grateful for slices of insight and the amelioration that PIAs can bring through transparency and a commitment to democratic values.

Cambridge, Massachusetts Gary T. Marx