What's in a Concept? Some Reflections on the Complications and Complexities of Personal Information and Anonymity

Gary T. Marx*

THE TOPIC OF ANONYMITY is conceptually and practically challenging. Among reasons for this are the multiple elements across different levels of analysis, varied contexts, and the variety of goals and dimensions that cross-cut these; 2) conflicting rationales and values; and 3) contested and/or opposing social, cultural and political trends and counter-trends.

In order to better understand anonymity, this conceptual analysis article raises and suggests possible answers to the following questions: what are the major types of information that can be associated with anonymity? (nine are identified, such as location and attitudes); what do we mean by personal information (using a series of concentric circles, distinctions are drawn between individual, private, intimate, unique and core identification); what are some of the major factors affecting behaviour involving anonymity, and judgments of anonymity (*e.g.*, the structure of the communication and the type of activity involved); what are the major values that support or oppose anonymity? (*e.g.*, openness in communication vs. accountability); what trends and counter-trends encourage or discourage anonymity (*e.g.*, technologies that make the meaningless meaningful as against increased freedom of choice with respect to identity); what broader principles are relevant to public policy in the area (*e.g.*, informed consent and reciprocity); and what kinds of questions should be asked in setting policy (*e.g.*, clear statement of goals, awareness of unintended consequences).

L'ANONYMAT EST UN SUJET stimulant sur les plans conceptuels et pratiques. Il y a diverses raisons pour cela, notamment : 1) la multiplicité des éléments provenant d'analyses à divers paliers, de contextes variés ainsi que des objectifs et des perspectives qui se dégagent de tout cela; 2) de justifications et de valeurs en conflit; enfin 3) de tendances ou de contre-tendances sociales, culturelles et politiques contestées ou opposées.

Afin de bien comprendre l'anonymat, par le biais d'une analyse conceptuelle, cet article propose des réponses possibles aux questions suivantes : quels sont les principaux genres de renseignements que l'on peut associer à l'anonymat? (neuf genres sont mentionnés, y compris le lieu et les attitudes); qu'entend-on par renseignements personnels? (utilisation d'une série de traits circulaires concentriques pour distinguer l'individu, la vie privée, l'intimité, les identificateurs uniques et fondamentaux); quels sont certains des principaux facteurs qui influent sur le comportement et les jugements en matière d'anonymat? (p. ex. la structure de la communication et le genre d'activité en cause); quelles sont les principales valeurs en faveur ou contre l'anonymat? (p. ex. la transparence de la communication versus l'imputabilité); quelles tendances ou contre-tendances encouragent ou découragent l'anonymat? (p. ex. les technologies qui donnent un sens à ce qui était vide de sens versus la liberté de choix et le respect de l'identité); quels principes plus larges sont importants pour les fins de la politique publique dans ce domaine? (p. ex. le consentement éclairé et la réciprocité); enfin, quelles genres de questions devrait-on se poser en élaborant la politique? (p. ex. des énoncés clairs des objectifs recherchés et une prise de conscience des conséquences non désirées).

---

* Professor Emeritus, MIT. This article expands on a presentation to "The Concealed I: Anonymity, Identity and the Prospect of Privacy" conference, University of Ottawa, March 2005.

What's in a Concept? Some Reflections on the Complications and Complexities of Personal Information and Anonymity

Gary T. Marx

It's a remarkable piece of apparatus.

–Franz Kafka, "In The Penal Colony"

"You ought to have some papers to show who you are." The police officer advised me.
"I do not need any paper. I know who I am," I said.
"Maybe so. Other people are also interested in knowing who you are."

–Bruno Traven, *The Death Ship*

1. INTRODUCTION

IN SPITE OF THE OFTEN WELL-INTENTIONED public declarations of those worried that anonymity is vanishing under the weight of new information-*invasive* technologies and rules *or* that its opposite, identifiability, is vanishing under the weight of new information-*protective* technologies and laws, the situation is much more complicated and complex than those concerns suggest. It does not lend itself to sweeping conclusions, however emotionally and ideologically satisfying these may be to their proponents. We all, of course, would like to be on the side of the angels. But in the case of anonymity the passionately offered didactic views of the prophets (whether of doom or utopia) need to be tempered with views of the concept wrestlers and empiricists.

As with many impassioned public controversies, answers to questions such as, "are things getting better or worse?" and "are current developments or proposals good or bad?" are both "yes" and "no" and "it depends." An important part of the scholar's job is to indicate just what judgments (whether of what is happening empirically, or of whether it is viewed as good or bad) *should*, and *do*, depend on.

I have studied these issues as part of an interest in the social impacts of information technology, particularly as these involve questions of privacy, anonymity, confidentiality, identity, civil liberties, surveillance, crime, deviance, and social control.[1] From my empirical inquiries, I here offer a bare-bones conceptual framework. Many additional empirical examples can be found in the articles in note 1 and at <www.garymarx.net>.

In the sections that follow, I argue that the topic is challenging and offers no easy answers because of 1) multiple elements across different levels of analysis, varied contexts, and the variety of goals and dimensions that cross-cut these; 2) conflicting rationales and values; and 3) contested and/or opposing social, cultural and political trends and counter-trends. I then suggest some principles and questions relevant to policy that can help us wend our way through the complexity.

<div align="center">*</div>

<div align="center">2. TYPES AND CONTEXTS OF ANONYMITY</div>

TO BEGIN, LET US COMPLICATE the question: "anonymous with respect to *what*?" The usual dictionary definition for anonymous is "not named or identified." Thus, the issue more broadly involves the availability or unavailability of a variety of kinds of information that may be known or identified about persons.

A major component involves answers to the "who are you?" question. An important factor conditioning evaluations of surveillance is the actual content, kind, or form of data gathered. Nine descriptive types of information about individuals which may be revealed or concealed can be noted. Before turning to these (detailed in Table 1), let me briefly consider four other analytic categories that also seem particularly relevant to understanding and judging surveillance.

The first involves *the inherent characteristics of the means* used (*e.g.*, bodily invasive, extends the senses, covert, degree of validity).[2] The second is the actual *application of the means* including the collection of the data and its subsequent treatment: Is the procedure competently and fairly applied and,

---

1. See generally Gary T. Marx, "Varieties of Personal Information as Influences on Attitudes Toward Surveillance" in Richard Ericson & Kevin Haggerty, eds., *The New Politics of Surveillance and Visibility* (Toronto: University of Toronto Press, 2006) 79-110, <http://web.mit.edu/gtmarx/www/vancouver.html>; Gary T. Marx, "Some Conceptual Issues in the Study of Borders and Surveillance" in E. Zureik & M. Salter, *Who and What Goes There? Global Policing and Surveillance* (Portland: Wilan, 2006) 11–35, <http://web.mit.edu/gtmarx/www/borders2.html>; Gary T. Marx, "Seeing Hazily (But Not Darkly) Through the Lens: Some Recent Empirical Studies of Surveillance Technologies" (2005) 30:2 Law & Social Inquiry 339, <http://web.mit.edu/gtmarx/www/hazily.html>; Gary T. Marx, "What's New About the New Surveillance? Classifying for Change and Continuity" (2004) 17:1 Knowledge, Technology and Policy 18 [Marx, "What's New"]; Gary T. Marx, "Murky Conceptual Waters: The Public and the Private" (2001) 3:3 Ethics and Information Technology 157, <http://web.mit.edu/gtmarx/www/murkypublicandprivate.html> [Marx, "Murky Conceptual Waters"]; Gary T. Marx, "What's in a Name? Some Reflections on the Sociology of Anonymity" (1999) 15:2 The Information Society 99, <http://web.mit.edu/gtmarx/www/ anon.html>; Gary T. Marx, "Ethics for the New Surveillance" (1998) 14:3 The Information Society 171, <http://web.mit.edu/gtmarx/www/ncolin5.html> [Marx, "Ethics"]; Gary T. Marx, *Undercover: Police Surveillance in America* (Berkeley: University of California Press, 1988).
2. See Marx, "What's New," *supra* note 1.

once collected, are trust and confidentiality sustained? Is there adequate security? Are undesirable consequences minimized or otherwise mediated? Traditional data-protection principles primarily apply to the second part of this. The third factor considers the *legitimacy* and *nature of the goals* that the surveillance tool is used for (*e.g.*, for the protection of health as against voyeurism). The fourth factor involves the *structure of the setting* in which the surveillance is used (*e.g.*, reciprocal vs. non-reciprocal, familial vs. non-familial).

       I have identified nine types of descriptive information on individuals:

1. Individual identification [the *who are you* question]
2. Shared identification [the *typification* question]
3. Geographical/Locational [the *where,* and beyond geography, *how to reach* question]
4. Temporal [the *when* question]
5. Networks and relationships [the *who else* question]
6. Objects [the *whose is it* question]
7. Behavioural [the *what happened* question]
8. Beliefs, attitudes, emotions [the *inner or backstage and presumed "real" person* question]
9. Measurement characterizations (predictions, potentials) [the *kind of person* question*, predict your future* question]

Space precludes a full consideration of each of these types. But, as an illustration of their complexity, and the need to go beyond common sense terms, let us consider several of them. I begin with the *who are you* question, as this involves issues of personal and private information relative to identity.

       Private information can refer either to the empirical status of information as known or not known (*e.g.*, President Clinton's affair was private but became public) or to norms which define information as being restricted (AIDS status is presumed to be private—as in being subject to the discretionary control of the person, even if it becomes widely known).

       Information that refers to an individual can be personal or impersonal. Any information that can be tagged to an individual is, in one sense, personal, but not all personal information involves expectations of privacy, nor, when it does, is this to the same degree.

       There is no easy answer to the questions "what is personal information?" and how does it connect to perceived assaults on our sense of dignity, respect for the individual, privacy, and intimacy. However, even giving due consideration to the contextual basis of meaning and avoiding the shoals of relativism as well as reification, it is possible to talk of information as being more or less personal. There is a cultural patterning to behaviours and judgments about kinds of personal information.

       These issues tie to broader sociology-of-information questions regarding norms about concealing and revealing information.[3] Here, violations may occur on the part of both the surveillance agent and the person of interest, in either failing to collect or to offer information, as well as in taking or offering

---

3. See Marx, "Murky Conceptual Waters," *supra* note 1.

information when it is not appropriate.[4] It is an interesting sociology-of-knowledge question why most academics and activists emphasize the involuntary collecting of personal information by authorities and organizations rather than their failing to gather information (*e.g.*, inquiring about arrest history for persons working with children) or the failure to reveal an outright deception on the part of individuals when they have an obligation to tell (*e.g.*, not informing a partner of a sexually transmitted or other disease, or not informing a potential buyer that a house for sale has a leaky roof) or inappropriately offering information as with public nudity, loud music or the revelation of intimate life details to strangers.

Surveillance involving information that is at the core of the individual, and more "personal," is likely to be seen as more damaging than that involving more superficial matters.[5] But what is that core and what radiates from it? How does the kind of information involved connect to that core?

Scientific explanation and moral evaluation require understanding what personal information is and how assessments of its collection, representation, and communication vary. This discussion is organized around the concepts in Tables 1—3. Table 1 describes kinds of information that may be gathered. The concentric circles in Table 2 show ways of characterizing information as individual, private, intimate, and sensitive with a unique core identity at the centre. Table 3 is more analytical and identifies cross-cutting dimensions that can be used to unite seemingly diverse, or to separate seemingly similar, forms. This permits more systematic comparisons and some conclusions about how the nature of the information gathered by surveillance is likely to be viewed.

Given the complex, varied and changing nature of the realities we seek to understand, I approach the task of classification humbly and tentatively, and note some limitations.

The above categories might be further organized under some broader concept such as biography. The concepts could be more systematically related to each other. At some point, combining enough shared-identity elements may permit identification of a specific individual. Unexpressed attitudes and thoughts could be seen as subtypes of behaviour. These differ from overt observable actions and from physical attributes such as having red hair (even here presenting natural red hair may be seen to represent a choice if the individual is aware of being able to dye hair, shave the head, or cover it with a hat). There may be disagreements about what behaviour is—traveling from one location to another is clearly a kind of behaviour, but what of staying in one place? Is a non-move still a move? Does it depend on the actor's awareness of the inaction?

---

4.   Related issues can also be seen with respect to organizations and the information that they offer and fail to offer. The main concern in this paper, however, is with individuals relative to the behaviour of larger, more powerful organizations.

5.   Yet such information is not only personal. Implications for groups matter as well, even if rarely acknowledged or studied. See generally Oscar H. Gandy Jr, *The Panoptic Sort: A Political Economy of Personal Information* (Boulder: Westview, 1993); Priscilla M. Regan, *Legislating Privacy: Technology, Social Values, and Public Policy* (Chapel Hill: University of North Carolina Press, 1995); Sheri A. Alpert, "Protecting Medical Privacy: Challenges in the Age of Genetic Information" (2003) 59:2 Journal of Social Issues 301. Our language, which gives us ready concepts about the individual and privacy, leads us away from seeing the social and shared components and how groups may be harmed or helped.

Furthermore, the concepts are not mutually exclusive. A given type of information about the person can fit into more than one category, either because the concepts deal with different aspects such as the body, time, place, relationships or behaviour, or because the information has mixed elements. For example, voice print, handwriting, and gait analysis are biometric and behavioural, in contrast to a form such as DNA, which is entirely biometric.

<div align="center">*</div>

<div align="center">3. INFORMATION ABOUT PERSONS</div>

WHAT IS PERSONAL INFORMATION? One approach defines it in broad terms as any information over which the individual has certain personal control or decisional rights.[6] Thus, facial image, copyrighted material, or the contents of one's medicine cabinet are personal, partly because they "belong" to the individual.

With new technologies, property ownership issues for information are unsettled. The lines between mine and not mine, self and other, copy and original, and the multiple meanings of public and private are increasingly blurred. Current digital rights management (DRM) controversies, for example, are about whether property that one "owns" in the form of DVDs or software can be altered at will, or must not be changed or copied without permission.[7]

Related controversies would likely follow the introduction of technologies for intercepting and reading brain waves and making sense of scents. In contrast to the "personal" property that one purchases, these are personal in their emanation from the individual. They may be viewed as an element of the person even after leaving the body. Spoken and written words share something with these forms, but will often be seen as less personal because willingly communicated. They involve assumptions about privacy, selfhood, and dignity.[8]

However, a property-control definition is not sufficient in that once control has been given up, personal information is still present. Thus, in the discarding of a pill container or a magazine subscribed to, personal information usually remains. In public settings, others may generally record the image and sounds the individual gives off.[9] These recordings do not cease to reflect the person as a result, even if they are re-creations and not "really" the person.

There are many situations where others have a right to access, use, and even "own" another's personal information. Yet the fact of their control does not then make it their "personal information" in the sense implied here. The ambiguity of language is a poor ally in this case.

---

6. See Alan F. Westin, *Privacy and Freedom* (New York: Atheneum, 1967); Charles Fried, "Privacy" (1968) 77:3 Yale Law Journal 475 [Fried, "Privacy"].

7. Julie E. Cohen, "DRM and Privacy" (2003) 18:2 Berkeley Technology Law Journal 575, <http://www.law.berkeley.edu/journals/btlj/articles/vol18/Cohen.stripped.pdf>.

8. See Edward J. Bloustein, "Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser" (1964) 39:6 New York University Law Review 962; Fried, "Privacy," *supra* note 6.

9. If unobtrusive "personal awareness assistance devices" for recording and coding interactions become common, the use of the devices will likely become an issue, and expectations about what can be appropriately gathered in "public" may change. Note that there is local legislation directed against certain forms of video surveillance in public.

Or, consider a person's DNA obtained from dental floss discarded as garbage. A California court has held that an ersatz garbage collector (as part of a lawsuit to prove paternity) could have DNA analysis performed on it.[10] Yet, the DNA still distinctively reflected the discarder. The "personal" remnants, of course, go beyond information and who is in "control." Note the presence of an intangible essence of the other person that is perceived to remain by those who purchase (often for outrageous sums) the costumes of movie stars and original manuscripts and letters of the famous.

Telephone numbers are another issue raising unresolved control questions. This became clear in the United States with the controversy over Caller ID in the late 1980s. The telephone companies simply offered the service as a *fait accompli*, essentially selling subscribers' numbers without their permission.

The act of *paying* for phone service (and paying even more for an unlisted number) would seem to imply control in a property rights sense over the number. Yet as legislation and regulations generally imply, the phone number is rented and "belongs" to the phone company (at least that is the case for land lines).

In the United States, a 2004 Federal Communications Commission ruling allows cell phone customers to retain their numbers if they change carriers within a calling area.[11] In Japan, phone numbers can be kept for life. Yet, the very property definition which permits individuals to "keep" the same number might also aid in making them subject to mass communication. It is easy to imagine a society in which all persons are given a never-to-be-changed telecommunications number at birth and, perhaps as the technology evolves, implanted with location and communication devices. The number would in one sense "belong" to the person, but would also bring costs, being the functional equivalent of a national ID card, and making the individual forever reachable, whether by friends, commercial and business interests, or the state.

It is curious that an individual has to pay for an unlisted number, while businesses pay to have their phone numbers listed. With respect to listed numbers, one can argue that the phone company should pay the individual consumers for being able to market their numbers in directories or deliver their number as part of a Caller ID service.

The question of formal ownership is distinct from the conditions of use and whether the number can be released if the phone company so chooses. However, its potential for reaching the individual and for probabilistic geodemographic analysis brings a personal component to it (*e.g.*, the social information revealed by telephone prefix or zip code—consider Harlem or Beverly Hills).

---

10. In the related case of *Moore v. Regents of the University of California*, 793 P.2d 479 (Cal Sup Ct 1990), a patient sued over what he claimed was the appropriation of his cells for commercial use, even though what was at stake was data based on his cells, not his literal cell material. The court found against him, since it claimed that a copy—rather than the original material—was involved.

11. See <http:// wireless.fcc.gov/wlnp/>.

As the above cases suggest, control is an important dimension and has policy implications.[12] Yet, something beyond control or possession is involved in defining personal information. We need a broader conception.

Another approach is to view any datum attached to a corporeal individual (identified by distinguishing characteristics of varying degrees of specificity) as "personal" because it corresponds to a person (*e.g.*, being identified as a citizen of the United States, a watcher of the Super Bowl, a middle-aged person, or an owner of an SUV). But information about an individual is not necessarily equivalent to *personal* information in a more restricted sense.

Knowledge about the kind of car one drives, when millions of people drive a similar car, is a pale form of "personal" information. It is more like impersonal information, although at a general level it serves to differentiate owners from non-owners and may convey symbolic meaning (*e.g.*, whether one owns a red convertible sports car or a black mega-pickup truck with flames painted on the front may be seen to be making a statement, whether intended or not, about the owner).

When we refuse, or resent, having information about ourselves taken, it is often because it is seen as "personal" or "none of your business."[13] In the United States, Prosser has identified four torts[14] and various revisions have been suggested;[15] these are causes of action for which suits can be brought in civil court by persons who feel that their privacy has been violated. Among the standard grounds are intrusion, disclosure, appropriation, and false light.

Several threads run through these. One is, of course, that the information is not "public," as in being immediately known to any casual observer (the way that the face or voice is). A second has to do with the nature of the information itself. It goes beyond many of the kinds of general information that can be associated with an individual. Private personal information needs to be located within the larger category of individual information.

Even within the more limited category, all private information (defined as information that is not automatically known about the other and that is subject to the actor's discretion to reveal it or give permission for it to be revealed) is not the same. Some goes to the very centre of one's person, while

---

12. Some analysts argue that the right to control the commercial use of one's image or of copyrighted material could be applied to other kinds of information associated with the individual, such as secondary uses of information on purchases. See Kenneth C. Laudon, "Markets and Privacy" (1996) 39:9 Communications of the ACM 92; James Rule and Lawrence Hunter "Towards Property Rights in Personal Data" in Colin J. Bennett & Rebecca Gran, eds., *Visions of Privacy* (Toronto: University of Toronto Press, 1999) 168–181.

13. Conversely, at other times it is the very impersonality of the treatment received that engenders resentment—the treating of (or communications to) the individual in a non-distinctive ("universalistic") fashion. This dynamic makes life interesting and also works against normative consistency, or at least general standards.

14. William L. Prosser, "Privacy" (1960) 48:3 California Law Review 338 at p. 389.

15. American Law Institute, *Restatement (Second) of Torts* (St Paul: American Law Institute Publishers, 1977) at ss. 652A–652I.

other information is peripheral or trivial (*e.g.*, sexual preference versus city of birth). Considerations of the private and personal involve looking at both content and procedure (both the means and the form).

Thus, some forms of communication (whatever their content) are viewed as personal: a sealed letter, a diary or a hushed conversation. This can be seen as a diagonal axis cutting across the types in Table 1. These types are defined by the presumed intent of the communicator. Their designation as a private form is distinct from whether or not we would view the content as personal and/or private. Contrast information about one's health status or one's favourite sports team sent via a post card with the same information sent in a sealed envelope. When there is unwarranted access to private means of communication, even with the most impersonal, general and inane of content, most persons would perceive a violation. Conversely, when private matters are inadvertently revealed in public, manners often require looking the other way or acting as if nothing has happened. The unexpected content can lead to a redefinition of the form.

Or, consider being secretly observed in a dressing room or taking a shower. What a hidden observer is likely to see will rarely be damaging to the observed person in any material or strategic sense. What generally matters is not what is seen (although scars, tattoos, or a colostomy are an exception), but the shock of being observed. This involves the means aspect and the violation of trust. However, for most of this discussion, I emphasize the content aspect.

<p align="center">*</p>

<p align="center">4. CONCENTRIC CIRCLES OF INFORMATION</p>

WE CAN THINK OF INFORMATION about persons as involving concentric circles of individual, private, and intimate and sensitive information. The outermost circle in Table 2 is that of *individual information*, which includes any data or category that can be attached to a person. The individual need not be personally known, or known by name and location, by those attaching the data. Individuals need not be aware of the data linked to their person.

Individual information varies from that which is relatively impersonal with minimal implications for an individual's uniqueness, such as being labelled as living in a flood zone or owning a four-door car, to that which is more personal, such as illness, sexual preference, religious beliefs, facial image, address, legal name and ancestry. The latter information has clearer implications for selfhood and for distinctly reflecting the individual.

The information may be provided by, or taken directly from, the person (*e.g.*, remote health sensors or a black box documenting driving behaviour). Or, it can be imposed onto persons by outsiders, as with the statistical risk categories of a composite nature used in extending credit.

The next circle refers to *private information* that is not automatically available. Absent special circumstances to compel disclosure, as with social security number for tax purposes or a subpoena or warrant for a search, such information is defined by discretionary norms regarding revelation. An unlisted

phone or credit card number and non-obvious or non-visible biographical and biological details are examples. As the case of "private parts" suggests, such information retains its moral, if not its existential, status as private, even when revealed, as with physical information at a nude beach. We can refer to information about the person that is not known by others and whose communication the individual can control as *existentially private.*

In contrast to such personal information of which the individual is aware is information with implications for life chances that is imposed from the outside of which the individual is often unaware. Much organizational categorization of individuals, of the kind that Foucault first called attention to, is encompassing, routine, invisible to the subject and artifactual.[16] This social sorting is fundamental to current social organization.[17]

Labelling by judicial, mental health, or commercial organizations may involve imputed identities (*e.g.*, a recidivism risk category) of which the individual is unaware (whether of the existence of the information or its content). Such imposed classifications are better seen as organizational secrets than as private information. The information may be considered personal and even sensitive where it is known by the individual.

Another distinction is whether, once known, an organization's information corresponds to how persons see themselves. This raises fascinating questions involving the politics of labelling and measurement validity and the consequences can extend even beyond the grave. In the case of the deceased, for example, note the controversy over the Church of Jesus Christ of Latter-day Saints' practice of posthumously baptizing some Jewish Holocaust victims.[18]

The disparity between technical labelling and self-definition may increase and become more contested as abstract measurements that are claimed to characterize the person and predict future behaviour based on comparisons to large databases become more prominent. And, apart from any organizational behemoths, self-definitions may be unclear as well. Note the uncertainty in the case of a female-to-male transsexual that underwent surgery so his outside would "match" his inside, who reports he is less sure than ever of what it means to be a woman or a man.[19]

The electronic and other spoors left by, and taken from, the individual, in connection with aggregate data about the behaviour of categories of

16. See Michel Foucault, *Discipline and Punish: The Birth of the Prison*, trans. A. Sheridan (New York: Vintage, 1979). Of course, the mere fact of artifactuality ought not to be automatically disqualifying. Many social artifacts have a persuasive logical and empirical validity. The trick is agreeing on and measuring the standards. Those in the business of educating others to the social construction of reality do a grave disservice when they treat all constructions as morally and pragmatically equal.

17. See David Lyon, ed., *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination* (London: Routledge, 2003); Geoffrey C. Bowker & Susan Leigh Star, *Sorting Things Out: Classification and its Consequences* (Cambridge: MIT Press, 1999).

18. Associated Press, "Mormons Said To Flout Vow Not to Baptize Deceased Jews" *New York Times* (11 April 2004) s. 1, p. 27.

19. Amy Bloom, *Normal: Transsexual CEOs, Crossdressing Cops, and Hermaphrodites with Attitude* (New York: Random House, 2003).

individuals, suggests an alternative simulated version of the person. In Germany this is referred to as a "shadow self." Poster writes of a "data double" reduced to pure information.[20]

Even when there is no disparity, such labelling serves as a new source of identity (*e.g.*, as a high SAT scorer or a low-cholesterol person). Organizational labelling has also become a marketable commodity—among many other forms, note the selling of background and credit rating scores at Sam's Club stores.[21]

The next circle is that of *intimate* and/or *sensitive* information. Intimate comes from the Latin *intimus*, meaning inmost. Used as a verb, the word intimate means to state or make known, implying that the information is not routinely known. Several forms of intimate or sensitive information can be noted.

Some "very personal" attitudes, conditions, and behaviours take their significance from the fact that they are a kind of currency of intimacy selectively revealed only to those we trust and feel close to. Such personal information is not usually willingly offered to outsiders, excluding exhibitionists and those seen to be lacking in manners. The point is not that the behaviour that might be observed is personal in the sense of necessarily being unique (*e.g.*, sexual relations are rarely highly individualizing, unknown or innovative), but that control over access affirms respect for the person and sustains the value of intimacy and the relationship. Persons who prematurely reveal their hole cards or private parts are likely to do poorly at both cards and love. As economists can often prove, in our culture, scarcity of desired resources tends to be associated with value.

We can also differentiate an intimate *relationship* from certain forms of information or behaviour that can be intimate—independent of interaction with others. E.M. Forster captures this in noting that we "radiate something curiously intimate when we believe ourselves to be alone."[22] This suggests a related form of intimacy—protection from intrusions into solitude or apartness. Whether alone or with trusted others, this protection implies a sense of security, of not being vulnerable, of being able to let one's guard down, which may permit both feelings of safety and of being able to be "one's self."[23]

This apartness, when protected by physical structures and manners (*e.g.*, bathroom activities) from others' observation, generally protects not against strategic disadvantage or stigmatization, but rather sustains respect for personhood.

The privacy tort of intrusion attempts to deal with the subjective and emotional aspects of harm resulting from incursions into solitude when personal borders and space are wrongly crossed. These are harder to define than harms

---

20. Mark Poster, *The Mode of Information Poststructuralisms and Social Context* (Chicago: University of Chicago Press, 1990).

21. According to Associated Press (9 March 2004), background screening software (US$39.77) is now being sold as a consumer product and a self-check service will be available soon. Such checks (for a fee) are said to give workers the opportunity to spot and correct problems in their personal records, as well as helping employers. Information brokers often stress that they are simply conveyor belts and not liable for mistakes. Some might see more than a little chutzpah here in a company's profiting from taking an individual's information, not sharing profits, and then selling it back to the individual.

22. E. M. Forster, *Where Angels Fear to Tread* (1905), <http://www.gutenberg.org/etext/2948>, ch. 7.

23. In a literal sense this is a structure or form issue rather than a content issue. Yet there is ambiguity in talking about "kinds of information" since this may include one or both.

such as unauthorized commercial use of a person's data, false light publicity, or public disclosure of private facts. The latter involve actions taken by the other *after* possession of information and usually some type of publication, even if no more than a sign in a shop window. With intrusion, in contrast, the process itself is objectionable.

Consider also moments of vulnerability and embarrassment that are observable in public, for example, the expression of sadness in the face of tragedy—as with a mother who has just lost a child in a car accident. Here, manners and decency require disavowing, looking the other way, not staring, let alone taking and publishing a news photo of the individual's grief.

Some information is "sensitive," implying a different rationale on the actor's part for information control and the need for greater legal protections. This includes the stigma that would devalue individuals in others' eyes or subject them to discrimination and the need to protect strategic information that could be useful to an opponent in a conflict situation or to a victimizer.

Sartre captures the security aspect of being seen when he noted, "What I apprehend immediately when I hear the branches cracking behind me is not that there is someone there; it is that I am vulnerable; that I have a body which can be hurt; that I occupy a place and I cannot in any case escape from this space in which I am without defense; in short, I am seen."[24] Yet this fear requires awareness, as does the challenging of the danger, and this element is lacking when the surveillance is transparent or unseen, as is increasingly the case.

Various US laws recognize information about finances, health and children as sensitive. The European Union's data protection directive requires special protection for "sensitive data," which it defines as involving information on race and ethnicity, political, philosophical and religious beliefs, health, and sexual life.[25]

More broadly, a central theme of Erving Goffman's work is that the individual, in playing a role and in angling for advantage, presents a self to the outside world that may be at odds with what the individual actually feels, believes, or "is" in some objective sense.[26] Through manners and laws, for most purposes, modern society acknowledges the legitimacy of there being a person behind the mask.

*

5. UNIQUE AND CORE IDENTITY

TWO FINAL CIRCLES at the centre involve a person with various identity pegs. These (whether considered jointly or individually) engender a *unique identity* ("only you" as the song says[27]) in being attached to what Goffman refers to as

---

24. Jean-Paul Sartre, *Being and Nothingness* (New York: Washington Square Press, 1993) p. 247.

25. EC, *Council Directive 94/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and the free movement of such data*, [1995] O.J.L. 28, <http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML>.

26. Erving Goffman, *The Presentation of Self in Everyday Life* (Garden City: Doubleday, 1959).

27. Harry Connick, Jr, *Only You*, audio CD (Sony Music Canada, 2004) at track 8.

an "embodied" individual who is usually assumed to be alive, but need no longer be. Knowing a unique identity answers the basic question raised by the children's Sesame Street television program, "who is it?" The question assumes the point of view of an outside observer trying to be honest, since individuals may prevaricate, have fluid identities, or, in rare cases, not know "who" they are.[28] It is from, and to, this identity that many other sources of potential information are derived or attached (radiating outward as well as being added) to the person.

The elements that make up the individual's uniqueness are more personal than those that do not, and as the degree of distinctiveness increases so does the "personalness" of the information.

Traditionally, unique identity tended to be synonymous with a *core identity* based on biological ancestry and family embedment. Excluding physically joined twins, each individual is unique in being the offspring of particular biological parents, with birth at a particular place and time. Parents and place of birth, of course, may be shared. Yet even for identical twins, if we add time of birth to the equation, the laws of physics and biology generate a unique *core identity* for each individual in the conjunction of parents, place and time of birth. This may of course be muddied by unknown sperm and egg donors, abandonment, and adoption.

For most people throughout history, discovering one's identity was not an issue. In small scale societies, where there was little geographical or social mobility and people were rooted in very local networks of family and kin, individuals tended to be personally known. Physical and cultural appearance and location answered the "who is it?" question.

Names may have offered additional information about the person's relationships, occupation, or residence (*e.g.*, Josephson, Carpenter, Frankfurt). Names are still sometimes presumed to offer clues to ethnicity, nationality, and class origin, and first names usually reflect gender. Titles such as Mrs or Dr convey additional information. Names popular in one time period that go out of fashion may also offer unintentional clues to approximate age.

However, the literal substantive information offered by a name is of little use when the observer, neither personally knowing, nor knowing about, the individual in question, needs to verify the link between the name offered and the person claiming it.

With large scale societies and the increased mobility associated with urbanization and industrialization, core identity came to be determined by full name and reliance on proxy forms such as a birth certificate, passport, national identity card, and driver's licence.[29] Yet, given adoption of children, the ease of

---

28.   Erving Goffman, *supra* note 26, in stressing the situationally specific nature of identity, would likely reject the idea of a core identity. However, my discussion begins with the objective facts of birth, not the individual's perception or social offerings regarding this.

29.   See Jane Caplan & John Torpey, eds., *Documenting Individual Identity: The Development of State Practices in the Modern World* (Princeton: Princeton University Press, 2001).

legally changing or using fraudulent names in the United States, and widely shared common names,[30] name may not be an adequate indicator of core identity. Nor, given technologies for forgery and the theft of identification, is the mere possession of identity documents sufficient for determining this.

The conventional paper forms of identification have been supplemented by forms more inherent in the physical person such as facial appearance or DNA, though even here we must remember that such measurements offer a representation of something inherent. The form of representation reflects choices rather than anything "given" in nature.

These measurements and transformations are forms of simulacrum.[31] According to the *Concise Oxford English Dictionary*, this can be a neutral "image of something" as well as a darker "shadowy likeness," "deceptive substitute," or "mere pretense."[32] Among the meanings of "simulate" are both "imitate" and "counterfeit." That contrast, of course, is what the fuss is all about: just how far in distorting the richness of the empirical should a simulation or a symbol go before it is rejected as invalid, inauthentic, inefficient, or ineffective?

With the expansion of biometric technology, a variety of indicators presumed to be unique (and harder to fake) are increasingly used (*e.g.*, beyond improved fingerprinting, we see identification efforts based on DNA, voice, retina, iris, wrist veins, hand geometry, facial appearance, scent, and even gait). The ease with which data collectors can secretly gather data that subjects involuntarily provide no doubt increases the appeal of these means.

The validity of these measurements as of 2005 varied significantly, from very high for DNA and fingerprinting (if done properly) to relatively low for facial recognition. There also are many ways of thwarting the surveillance. Even when validity is not an issue, biological indicators are not automatic reflections of core identity—although they may offer advantages such as being ever present and never forgotten, lost or stolen. To be used for identification there must be a record of a previously identified person to match to the indicator.[33] These need not lead to literal identification of a person, but rather to answering whether the material presumed to reflect a unique person is the same as that in a database to which it is compared (*i.e.*, "is this the same person?" *whoever* it is). With data from multiple events, police may know because of matching that the same person is responsible for crimes but not know who the person is.

---

30. With respect to accuracy and currency, note multiple listings for the same person as a result of the inclusion of former or several residences, failure to purge the deceased, or the exclusion of those with unlisted numbers and those using pseudonyms. Marc Pairman the lead singer of a group called "Gary Marx and the Sisters of Mercy" calls himself Gary Marx. It is also possible that some of those using the name Marc Pairman might not really be him. ☺

31. See Jean Baudrillard, *Selected Writings*, ed. Mark Poster (Stanford: Stanford University Press, 1988); William Bogard, *The Simulation of Surveillance: Hypercontrol in Telematic Societies* (New York: Cambridge University Press, 1996).

32. *The Concise Oxford English Dictionary*, *s.v.* "simulacrum" (New York: Oxford University Press, 2004).

33. While one must be very wary of any single determinism, this is a nice example of a social process in which a technical development drives a social development. Note the emergence of law enforcement, military and other DNA and fingerprint and voice print databases. For effective use, the unidentified DNA (and other) data "require" organizational databases tied to known individuals. Here we see an example of *surveillance creep* or, and often, *gallup*. Of course where there is an accessible likely suspect, a broad database is not required. Yet legal and moral authority and logistic accessibility are still required to carry out the match. An alternative is trickery. The ingenious means of discovery employed by those involved in paternity suits or who otherwise seek to use DNA to prove something is the stuff of the six o'clock news and murder mysteries. These means vary from going through the garbage to planting spent condoms.

Police files are filled with DNA and fingerprint data that are not connected to a core legal identity. Some jurisdictions such as New York have "John Doe" programs in which charges can be filed based on DNA profiles alone, even though there is as yet no name to which to link the information. This is intended to permit prosecution should such a link be established.[34]

The question "who is it?" may be answered in a variety of other ways that need not trace back to a biologically defined ancestral core or legal name. For many contemporary settings, what matters is determining the presence of attributes warranting a certain kind of treatment or continuity of identity (is this the *same* person) or the location of the individual, not determining who the person "really" is as conventionally defined.

A central policy question is how much and what kind of identity information is necessary in various contexts. In particular, whether identification of a unique person is appropriate; and if it is, what form it should take. I have suggested distinctions between anonymity with respect to core biological identity (*e.g.*, an individual as unique because he or she is born at a particular place and time as a result of the biological uniting of two parents) as linked to a legal identity, as against various other pseudonymous forms of identity such as a national identification number. Pseudonyms may or may not be linked to a given name or location. Sometimes what matters is being able to locate an individual or authenticate some aspect of his identity, rather than literally knowing the core biological/legal identity. This may be for purposes of communication or denying or granting some form of access or privilege.

Table 3 offers a different, more abstract approach in considering factors that may cut through the various descriptive forms in Table 1, uniting the seemingly dissimilar and separating the seemingly similar. With respect to these categories, the absence of anonymity and the involuntary revelation of personal information become more problematic the more that the values on the left side of the table are present.

I hypothesize that, other factors being equal, when anonymity or non-revelation are appropriate but are not honoured, there is an additive effect: the more that the values on the left side of the table are present, the greater is the perceived wrong in the collection of personal information. The worst possible cases involve a core identity, a locatable person, information being attached to the person that is personal, intimate, sensitive, stigmatizing, strategically valuable, extensive, biological, naturalistic, predictive and revealing of deception, and an enduring and unalterable documentary record.

---

34.   See William K. Rashbaum, "New York Pursues Old Cases of Rape Based Just on DNA" *New York Times* (5 August 2003) A1.

*

6. ONE SIZE DOES NOT FIT ALL

BEYOND THE ABOVE TYPES and dimensions of personal information, variation in behaviour and judgments will vary depending on:

1. *types of communicator/recipient* (*e.g.*, children and other dependants, responsible and irresponsible adults, law enforcers, persons vulnerable to retribution for reporting wrong-doing, those seeking information versus those from whom information is sought, sending information/communication versus receiving it)
2. *the structure of communication* (*e.g.*, one-on-one, one-to-many, many-to-one and reciprocal or non-reciprocal, real or stale time, moderated and unmoderated groups)
3. *types of activity* (*e.g,* browsing, requesting information, posting bulletin boards, email, discussion groups)
4. *content/goals* (*e.g.*, games, self-help groups, hotlines, commerce, politics, science, protecting the sender of a communication or the recipient)
5. *the national and cultural borders* that communication invisibly crosses and the types of response (*e.g.*, prohibited, required, optional but favoured or disfavoured, laws, policies, manners). Even if one could agree on a policy regarding computer-related anonymity, there is no central world internet authority to implement it and doing this would be technically difficult.

*

7. RATIONALES FOR AND AGAINST ANONYMITY

LET US NEXT CONSIDER some values supporting and opposing anonymity in communication and some broader value conflicts within which these fit.

The public policy questions raised by technologies for collecting personal information are more controversial than many other issues such as ending poverty and disease. In those cases, the conflict involves asking "how" rather than "why." The questions raised by the concealment and revelation of personal information are like some relationships in which persons can not live with each other, but neither can they live apart. The issue becomes under what conditions do they co-exist? So it is with anonymity and identifiability. There are existential dilemmas and in many cases we are sentenced to a life of tradeoffs.

I often ask my students what society would be like if there were absolute transparency and no individual control over personal information—if everything that could be known about a person was available to anyone who wanted to know. Conversely, I also ask them what society would be like if there were absolute opaqueness such that nothing could be known about anyone except what they chose to reveal. The absolute anonymity versus absolute identifiability is a strand of this. Both of course would be impossible and equally unliveable, but for different reasons. To have to choose between repression and anarchy is hardly a choice between a pillow and a soft place.

The hopeful Enlightenment notion that with knowledge problems will be solved holds more clearly for certain classes of physical and natural science questions than for many social questions. Certainly those who live by the pursuit of truth dare not rain on that parade. Yet there is a difference between knowledge as providing answers, as against wisdom. Current debates over anonymity and identifiability in electronic communications would greatly benefit if better data were available, but the issue would not disappear because the value conflicts and varied social and psychological pressures remain.

A cartoon image nicely captures this—we see a tanker truck carrying hazardous material with a sign on the back which says, "The scientific community is divided. Some say this stuff is dangerous, some say it isn't."[35] So it is with this issue. The divisions do not reflect ignorance, stupidity, ill-will and evil on one side and empirical truth, wisdom, benevolence and righteousness on the other. Rather they reflect varying degrees of empirical truth on both sides and differing value priorities. Being able to disentangle these is vital for our understanding and for developing policy. Let us consider the values and goals question.

Among the most common justifications for full or partial anonymity are:

1. to facilitate the flow of information and communication on public issues
2. to obtain personal information for research in which persons are assumed not to want to give publicly attributable answers or data
3. to encourage attention to the content of a message or behaviour, rather than to the nominal characteristics of the messenger which may detract from that
4. to encourage reporting, information seeking, communicating, sharing and self-help for conditions that are stigmatizing and/or which can put the person at a strategic disadvantage or are simply very personal
5. to obtain a resource or encourage a condition using means that involve illegality or are morally questionable, but in which the goal sought is seen as the lesser evil
6. to protect donors of a resource, or those taking action seen as necessary but unpopular, from subsequent obligations, demands, labelling, entanglements or retribution
7. to protect strategic economic interests, whether as a buyer or a seller
8. to protect one's time, space and person from unwanted intrusions
9. to increase the likelihood that judgments and decision-making will be carried out according to designated standards and not personal characteristics deemed to be irrelevant. A well known cartoon of two computer literate dogs captures this, as one says to the other, "on the internet no one knows you're a dog."[36]

---

35. Mischa Richter, "The scientific community is divided: some say this stuff is dangerous, some say it isn't," cartoon, *The New Yorker* (21 March 1988), <http://www.cartoonbank.com/assets/1/22815_m.gif>.
36. Peter Steiner, "On the Internet no one knows you're a dog," cartoon, *The New Yorker* (5 July 1993), <http://www.cartoonbank.com/assets/1/22230_m.gif>.

10. to protect reputation and assets
11. to avoid persecution
12. to enhance rituals, games, play and celebrations
13. to encourage experimentation and risk taking without facing large consequences, risk of failure or embarrassment
14. to protect personhood or "it's none of your business"
15. traditional expectations

A consideration of contexts and rationales where anonymity is permitted or required must be balanced by a consideration of the opposite. The rationales here seem simpler, clearer and less disputed. While there are buffers and degrees of identification, the majority of interactions of any significance or duration tilt toward identification of at least some form.

Central to many of the contexts where some form of identifiability is required we find the following rationales:

1. to aid in accountability
2. to judge reputation
3. to pay dues or receive just desserts
4. to aid efficiency and improve service
5. to determine bureaucratic eligibility—to vote, drive a car, fix the sink, cut hair, do surgery, work with children, collect benefits, enter or exit (whether national borders, bars or adult cinemas)
6. to guarantee interactions that are distanced or mediated by time and space
7. to aid research (links to other types of personal data and longitudinal data)
8. to protect health and consumers
9. to aid in relationship building
10. to aid in social orientation

*

8. VALUE CONFLICTS

THE OFTEN CONFLICTING GOALS of anonymity and identifiability are nestled within a broader set of information and societal value conflicts: enduring value conflicts and ironic, conflicting needs and consequences which make it difficult to take broad and consistent positions regarding the revelation or concealment of personal information as this involves information technology.

For example, we value both the individual and the community. We want both liberty and order. We seek privacy and often anonymity, but we also know that secrecy can hide dastardly deeds and that visibility can bring accountability. But too much visibility may inhibit experimentation, creativity and risk taking. In our media-saturated societies we want to be seen and to see, yet also to be left alone. Note the desire to reveal as seen in popular talk shows and celebrity tell-all books and public relations activities.

We value freedom of expression and a free press but do not wish to see individuals defamed or harassed. We desire honesty in communication and also civility and diplomacy. We value the right to know, but also the right to control

personal information. The broad universalistic treatment citizens expect may conflict with the efficiency driven, specific treatment made possible by fine-honed personal surveillance data. The expectation that one should be judged as an individual and in context may conflict with the greater rationality and predictive success believed to be found in responding to aggregates.

Many discussions between those who look optimistically at information technology as the solution and those who view it as the problem reflect the Hindu tale about blind persons and the elephant, in which each observer offers a plausible identification for one part of the elephant (*e.g.*, the tail as rope).[37] That is, a legitimate goal or social trend is identified, but other confounding ones are ignored or denied.

*

9. SOME TRENDS AND COUNTER-TRENDS

A FINAL COMPLICATION working against quick conclusions involves trends and counter-trends. Let me locate changes in anonymity alongside of a number of other social developments that suggest issues for social research and which make broad and unitary social and moral assessments challenging. There are ironic trends and counter-trends and unintended consequences. The field is fluid with new opportunities and problems ever emerging. Solving one problem may create another in an endless dialectical dynamic.

The very rapid changes in data collection, storage, and analysis are central to the topic. Through 2003, processing speeds had doubled every eighteen months and storage capacities had doubled every year.[38]

Means of data analysis once restricted to governments and the largest organizations are available on a much wider scale to smaller organizations and individuals. Diverse kinds and sources of data are increasingly woven into a network. Computing is becoming ubiquitous and automated: sensors that passively read and send (with no action required on the part of the actor) remote signals to the internet and elsewhere are increasingly found in objects (*e.g.*, computing and communications devices, switches, groceries, cars, tools, weapons, clothes), persons, and environments (roads, walls, doors). Through a "value-added" model, the aggregation and analysis of data collected in varying formats and for varying purposes in turn creates new data and models. More information is also available for analysis because ever more is being kept rather than culled. It is now less expensive to store information than to discard it.

Given such changes, one might evoke Paul Simon's classic line in the song "Slip Slidin' Away" ("the nearer your destination the more you slip slidin' away").[39] Among trends that many see as worrisome and actually, or potentially, threatening traditional values of democratic societies are:

---

37. Paul Gladone, *The Blind Men and the Elephant: John Godfrey Saxe's Famous Indian Legend* (London: Egmont Children's Books, 1973).

38. National Research Council, *Engaging Privacy and Information Technology in a Digital Age: Issues and Insights* (Washington D.C.: The National Academies Press, 2007).

39. Simon and Garfunkel, *The Concert in Central Park*, audio CD (Warner Bros., 1982) at track 11.

1. *The decline of anonymity.* The ability to be unnoticed has declined significantly, although this is not the same as being uniquely known.
2. *Making the meaningless meaningful.* Once noticed, the ability to remain unidentified, whether by core identity, or some other specific measure has declined.
3. *Colonization of time, space and physical borders.* Whether voluntarily or involuntarily on the subject's part, the ability to discover and track the varied forms of individual information in real-time across physical barriers, locations, and over time has significantly increased.
4. *Increased validity, but still far from ideal for many purposes.* When correctly applied, current core identification technologies show a high degree of validity relative to the cruder bodily measurement and eye-witness techniques of the 19th century. These are better for the time period when they are applied than for the past or future.
5. *Category expansion.* There is a significant expansion of ways of measuring and classifying individuals and contexts, and these are retrospective, as well as prospective. These abstract characterizations that symbolize personal characteristics involve behaviour as well as presumed essence (whether physiological or moral). These often are, but need not necessarily be, attached to a core identity. These involve greater precision than traditional measures. We see composite measures that are increasingly removed from the "natural" relatively uncomplex factors that composed personal information prior to, and even during, industrialization.
6. *The merging of previously compartmentalized data.* The ability to be known about as a result of combining indicators has significantly increased.
7. *Apart from technical developments that permit involuntarily collecting personal information, there has been a major expansion of laws, policies and procedures mandating that individuals provide information.* Whether related to effectiveness, crises, or fairness, access to participation in modern life (voting, government benefits, employment, building or gated community access, etc.) increasingly requires some form of identity validation.
8. *The integration of life activities with the generation of personal data.* We increasingly live in ways that automatically provide personal information as part of the activity—*i.e.*, the use of credit cards, communication and driving.
9. *The blurring of lines between public and private places makes personal information more available.* Note the privatization of places traditionally seen as "public," such as shopping malls and industrial parks (with legal means of collecting personal information). Or consider the blurring of the lines between home and work and the merging of public and private databases and the ability of technologies to reveal some aspects of what is within a

private space without the need to literally enter it, *e.g.*, thermal imaging or cameras in pubic places that are aimed at private places. Other examples are the availability of web and related searches in finding and merging personal information that had been *de facto* private because of spatial and temporal separation and the presumed ability to learn about non-consenting individuals by generalizing from those sharing attributes who voluntarily provide the information (*e.g.,* focus groups).

These trends suggest the familiar "no where to run," tightening of the noose, decline of private space, privatization of public space, Leviathan all-knowing political, commercial, and even interpersonal State of the Dystopic Imagination. Yet, however powerful as an indicator of a social trend and as a raiser of consciousness, this view must be tempered by noting opposing developments.

The current situation is dynamic and rapidly changing. We might equally invoke the Beatles' claim that "things are getting better all the time"[40] in observing some opposing trends involving the ironic vulnerabilities of any system of control, as well as broader historical trends.[41] Among some counter-trends:

1. *Increased freedom of choice.* Individuals in some ways are freer both morally and tactically to make or remake themselves than ever before. Some identities that historically tended to be largely inherited such as social status or religion can more easily be changed. Other identities have become culturally more legitimate, such as divorce, homosexuality, birth out of wedlock, and adoption, with a subsequent decline in traditional stigmas and the need to be protective of certain kinds of information.[42] Even seemingly permanent physical attributes such as gender, height, body shape, or facial appearance can be altered, whether by hormones or surgery or beauty parlours. The ultimate change is the emerging technology of total face transplants. Television "make over" shows and self-altering products reflect related strands of this. These developments reflect the emergence of a more protean self and the self as a commodity and an object to be worked on, just as one would work on a plot of land or carve a block of wood. Identities in some ways are becoming relatively less unitary, homogeneous, fixed and enduring, as the modernist idea of being able to choose who we are continues to expand, along with globalization processes. This is aided by the expansion of non-face-to-face interaction.

---

40.   Beatles, *Sgt. Pepper's Lonely Hearts Club Band*, audio LP (Parlophone, 1967) at track 4.

41.   The careful listener, however, will hear this relativized, as they whisper, "they couldn't get much worse."

42.   Increased freedom of choice can exist with increased volume and intensity of surveillance. This is merely to suggest that the kinds of information individuals wish to keep private changes with social and cultural change, not that the overall amount we wish to conceal declines. That is an empirical question which must take into account the absolute amount there is to be known about a person, a factor that has markedly increased and continues to increase in recent centuries. New diagnostic means involving DNA and predictive profiles for at-risk individuals may create new forms of stigma. The case for a relative increase in surveillance is dependent on the ratios of what there is to be known of interest, what the technology is capable of, and the actual extent of its application.

2. Sex change operations are at one extreme. But more common are the new identities created through the increased intermarriage of ethnically, racially, religiously and nationally distinct groups. An increase in children of mixed marriages, those holding dual-citizenship, immigration, tourism and communities in cyberspace illustrate this. New categories for marginal, hybrid and anomalous groups will appear. As just one example take the millions of Americans who, as products of a mixed marriage, consider themselves *both* Christian and Jewish, White and Black, or Asian and Hispanic.

3. *New opportunity structures for exercising choice.* The distance mediated interaction of cyberspace which calls forth new means of authentication also opens up a vast potential for offering various forms of alternative or prevaricated individual information. Cyberspace as play (*e.g.*, internet service providers offering online aliases and fantasy chat rooms) encourages this, while also offering new protections for identity (although this can also lead to new forms of protection, as well as violation).

4. *New functional alternatives to core identity.* The absolute number and relative importance of non-core forms of identity offering varying degrees of anonymity has increased. There is a significant expansion in the variety of pseudonymous certification mechanisms intended to mask or mediate between the individual's name and location, yet still convey needed information. As more and more actions are remotely tracked in cyberspace (*e.g.*, phone communication, highway travel, consumer transactions), the pseudonym will become an increasingly common and accepted form of presenting the self for particular purposes (whether as a unique individual or as a member of a particular category). A cartoon showing a talking bird who speaks but only anonymously illustrates this.[43]

5. *Enhanced chances for neutralization.* Beyond the expansion of life style/identity choices we see the ironic emergence of markets for counter-surveillance offering a vast array of methods to protect individual information, whether by blocking, distorting, deceiving or destroying the surveillance means. Much of this represents a righteous response to the creeping or galloping expropriation of personal information, yet some also represents new opportunity structures for violation. The ease of presenting fraudulent identities divorced from the traditional constraints of localism and place and time is central to crimes of identity theft.

---

43. Mort Gerberg, "Yes, he does speak but only on condition of anonymity," cartoon, *The New Yorker* (10 July 2000), <http://www.cartoonbank.com/assets/1/44031_m.gif>.

6. *Significant improvements in technologies for protecting individual information.* With encryption there is the potential for a degree of confidentiality in communication, and enhanced accountability and data protection never before seen. Technologies and services for protecting personal information are increasingly available, from shredders to home security systems to various software and privacy protection services.

7. *New normative protections and awareness.* There has been a significant expansion of laws, policies and manners that limit and regulate the collection of personal information and its subsequent treatment. There has been some growth in choice and opt-in systems. This ties to the broader twentieth-century expansions of civil liberties and civil rights, as well as to particular crises. Whether these go far enough, are effective, and how they compare across institutions and cultures are important research questions.

The trends and counter-trends seen in the two broad perspectives offered above work against sweeping generalizations beyond this one against sweeping statements. Considered together some of the above developments are ironic and contradictory. I take this as a sign of reality's ability to overflow our either/or categories and the need to avoid simplistic theorizing, as well as the need for empirical research.

Let me move from the above considerations which reflect my views as an empirical scientist to some ideas more explicitly reflecting both data and values to inform policy.

Scholars are in the business of elevating considerations of complexity to a high art form, but others need to create policies and laws and make decisions about revealing or concealing personal information. We can endlessly debate these questions. But those setting policy must act. What can an academic analysis offer? Conceptual tools, empirical data, awareness, vigilance and self-reflection is one answer! We must be mindful of the cultural background assumptions (both empirical and normative) that, like icebergs, lurk beneath the surface of our taken-for-granted worlds. We can also offer ideas as buoys or channels to direct policy.

*

10. PRINCIPLES TO INFORM PUBLIC POLICY

WITH RESPECT TO QUESTIONS of ethics and policies for governing the collecting, storing, accessing, merging, analysing and communicating of personal information, the principles in Table 4 are central. Many of these were first expressed in the Code of Fair Information Practices developed in 1973 for the US Department of Health, Education & Welfare.[44] They are also now found in

---

44. US Department of Health, Education and Welfare, Secretary's Advisory Committee on Automated Personal Data Systems, *Records, Computers, and the Rights of Citizens* (July 1973), <http://www.aspe.hhs.gov/datacncl/1973privacy/tocprefacemembers.htm> [*Records, Computers*].

various European and Asian directives such as the privacy guidelines of the Organization for Economic Cooperation and Development (1981).[45]

The 1973 Code offered a *principle of informed consent* in which the data collection is not to be done in secret, and individuals are to be made aware of how it will be used, and where appropriate, consent to it; a principle of *inspection and correction* in which individuals are entitled to know what kind of information has been collected and to offer corrections and emendations; a *principle of data security* in which the information will be protected and precautions taken to prevent misuses of the data; a *principle of validity and reliability* in which organizations have a responsibility to insure the appropriateness of the means used and the accuracy of the data gathered; and *a principle of unitary usage* in which information gathered for one purpose is not to be used for another without consent.[46]

As new information technologies, uses, and problems have appeared, additional principles are necessary. From my research on new information technologies, it is clear that the above principles, however important, need to be expanded. I would thus suggest the following additional principles: a *sanctity of the individual and dignity principle* in which there are limits (even with consent) on the taking, volunteering and commodification of personal information; a *golden rule principle* in which those doing the surveillance would agree to be the subjects of information gathering under comparable circumstances; a *principle of consistency* such that broad ideals rather than the specific characteristics of a technology should govern surveillance practices; a *principle of morality* in which the fact that a tactic is legal is not sufficient justification for using it, apart from broader ethical considerations; *principles of relevance* and of *minimization* such that only information that is directly relevant and necessary for the task at hand is gathered (minimization refers to both the amount of information gathered and the intrusiveness/invasiveness of the means); a *principle of joint ownership of transactional data* such that both parties to a data creating transaction should agree to any subsequent use of the data, including the sharing of benefits if appropriate; broadening of the principle of informed consent to favouring opting-in over opting-out and a *principle of co-determination, or at least consultation* regarding policies; *a principle of restoration* such that in a communications monopoly context those altering the privacy status quo should bear the cost of restoring it; *a safety net or equity principle* such that a minimum threshold of information protection should be available to all; a *principle of equal treatment* such that surveillance deemed to be invasive, but appropriate, is applied to all members of an organization, not just the least powerful members; a *reciprocity or equivalence of tactics principle* in which in situations of legitimate conflict of interest all parties can use the same tactics; a *principle of timeliness* such that data are expected to be current and

---

45. Organisation for Economic Cooperation and Development (OECD), *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (Paris: OECD, 1980), <http://www.oecd.org/document/18/0,2340,en_2649_201185_1815186_1_1_1_1,00.html>.

46. *Records, Computers, supra* note 44.

information which is no longer timely should be destroyed; a *principle of the periodic review and evaluation of data collection policies* as broadly defined; a *principle of human review* such that an automated decision is always subject to review by a person; a *principle of redress* such that those subject to inappropriate surveillance or unfairly hurt by it have adequate mechanisms for discovering and being compensated for the harm; a *less worse alternative means principle* in which means are compared to each other; and a *sometimes it is better to do nothing principle* in which the consequences of inaction are compared to those of action.

I offer the above in a tentative spirit that demands discussion. Given the general, idyllic and antiseptic nature of such principles, they should be viewed as ideas to provoke thought and discussion, not as a unified body of theory, nor as principles to be rigidly applied or equally weighed. There would be widespread agreement on some, while others are more controversial, such as the idea of joint ownership of transactional data. Some of them may conflict or offer little help with complicated issues (*e.g.*, the third-party possession of records), and some cannot be easily implemented. They are best viewed as awareness-honing devices.

These principles can be stated in the form of the questions in Table 5 to be asked about policy development for a given area such as anonymity.[47] In general, the more these questions can be answered in a way consistent with the underlying principles, the more legitimate the surveillance is. Or conversely, the less they can be answered in this way, the less legitimate it is. Legitimacy in one area need not generalize to others (*e.g.*, a valid tactic may be poorly applied even where the goal is valid, or a valid tactic may be competently applied to an inappropriate goal). Valid means and goals can be found in settings where appropriate authorization procedures are absent or not followed.

Certainly, these principles and questions cannot be automatically transferred to situations such as those of public order and health, criminal investigations, national security, or times of sudden catastrophic natural or human-created disasters. A central point of much sociological analysis is to call attention to the contextual nature of behaviour. Yet, common sense and common decency argue for the consideration of these principles, absent compelling circumstances.

Whatever action is taken, there are likely to be costs, gains and trade-offs. At best, we can hope to find a compass—rather than a map—and a moving equilibrium—rather than a fixed point—for decision making.

This article opens with a tongue-in-cheek statement from Kafka's short story, "In the Penal Colony," in which a new technology is referred to as "a remarkable piece of apparatus."[48] This is a very sophisticated machine for punishing inmates, which was invented by a corrections officer. The story ends

---

47.   A fuller statement of this approach is in Marx, "Ethics," *supra* note 1.

48.   Franz Kafka, "In the Penal Colony" in *The Penal Colony*, trans. by Willa & Edwin Muir (NY: Schocken Books, 1968) 191.

with the officer dying after falling into the machine he had created.[49] While it is premature, and perhaps even sacrilegious, to conclude that information technology will destroy rather than save us, Frankensteinian outcomes are not always figments of the literary or psychoanalytic imagination. Research and vigilance, however, may work against this.

49. Note also Nathaniel Hawthorne, "The Birthmark" in R. P. Blackmur, ed., *The Celestial Railroad and Other Stories* (NY: Penguin Press, 1980) 203–220. In this story, first published in 1843, an alchemist, in seeking to successfully rid his wife of a small blemish, accidentally kills her. The operation was a success but the patient expired.

Table 1. Some types of descriptive information connectable to individuals

1. Individual identification
    [the "who" question]

    - Ancestry
    - Legal name
    - Alpha-Numeric
    - Biometric (natural, environmental)
    - Password
    - Aliases, nickname
    - Performance

2. Shared identification
    [the "typification-profiling" question]

    - Gender
    - Race/ethnicity/religion
    - Age
    - Education
    - Occupation
    - Employment
    - Wealth
    - DNA (most)
    - General physical characteristics (blood type, height) and health status
    - Organizational memberships
    - Folk characterizations by reputation—liar, cheat, brave, strong, weak, addictive personality

3. Geographical/Locational
    [the "where," and "beyond geography, how to reach" question]

    A. Fixed
        - Residence
        - Telephone number (land line)
        - Mail address
        - Cable TV

    B. Mobile
        - Email address
        - Cell phone
        - Vehicle and personal locators
        - Wireless computing

4. Temporal
    [the "when" question]

    - Date and time of activity

Table 1 (continued)

5. Networks and relationships
    [the "who else" question]

- Family members, married, or divorced
- Others the individual interacts/communicates with, roommates, friends, associates
- Others co-present (contiguous) at a given location (including in cyberspace) or activity including neighbours

6. Objects
    [the "which one," "whose is it," and "who used it" question]

- Vehicles
- Weapons
- Animals
- Communications device
- Contraband
- Land, buildings, and businesses

7. Behavioural
    [the "what happened" question]

- Communication
- Fact of using a given means (computer, phone, cable tv, diary, notes or library) to create, send, or receive information (mail covers, subscription lists, pen registers, email headers, cell phone, GPS)
- Content of that communication (eavesdropping, spyware, library use, book purchases)
- Economic behaviour—buying (including consumption patterns and preferences), selling, bank, credit card transactions
- Work monitoring
- Employment history
- Norm and conflict related behaviour—bankruptcies, tax liens, small claims and civil judgments, criminal records, suits filed

8. Beliefs, attitudes, emotions
    [the "inner or backstage and presumed 'real' person" question]

9. Measurement characterizations (past, present, predictions, potentials)
    [the "kind of person, predict your future" question]

- Credit ratings and limits
- Insurance ratings
- College readiness/acceptability scores
- Civil service scores
- Drug tests
- Truth telling
- Psychological tests and profiles
- Occupational placement tests
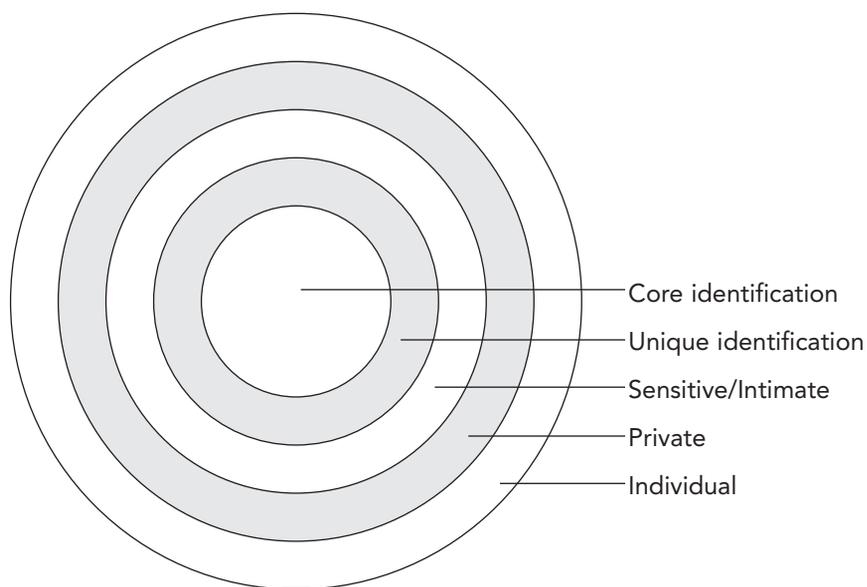- Medical

Table 2. Types of information on the embodied

Core identification
Unique identification
Sensitive/Intimate
Private
Individual

Table 3. Some dimensions to analyse types of descriptive individual
         information

1. personal	
• yes	• no (impersonal)
2. intimate	
• yes	• no
3. sensitive	
• yes	• no
4. unique identification	
• yes (distinctive but shared)	• no (anonymous)
• core	• non-core
5. locatable	
• yes	• no
6. stigmatizing (reflection on character of subject)	
• yes	• no
7. prestige enhancing	
• no	• yes
8. reveals deception (on part of object)	
• yes	• no
9. strategic disadvantage to subject	
• yes	• no
10. amount and variety of data	
• extensive, multiple kinds	• minimal, single kind
11. documentary (re-usable) record	
• yes [permanent?] record	• no
12. attached to or part of person	
• yes	• no
13. biological	
• yes	• no
14. naturalistic (reflects "reality" in obvious way, face validity)	
• yes	• no
15. information is predictive rather than reflecting empirically documentable past and present	
• yes	• no
16. shelf life	
• enduring	• transitory
17. alterable	
• yes	• no

Table 4. Some principles for guiding manners, policies, and law with respect to personal data

1. the sanctity and dignity of the individual principle

2. the principle of consistency

3. the principle of asking the golden rule question

4. the principle of moral acceptability over legal acceptability

5. the principle of the informed subject

6. the principle of the consenting subject

7. the principle of opting-in favoured over opting-out

8. the principle of co-determination, or at least consultation, in setting policies

9. the principle of relevance

10. the principle of minimization with respect to the kind and amount of data and degree of intrusion and invasion in collection and data essence

11. the principle of validity and reliability

12. the principle of timeliness

13. the principle of inspection and correction

14. the principle of data security

15. the principle of confidentiality

16. the principle of joint ownership of transactional data

17. the principle of unitary usage (non-data migration)

18. the principle of restoration

19. the principle of redress (including compensation for harm)

20. the principle of human review of automated decisions

21. the principle of periodic review and evaluation of policies

22. the safety net or equity principle

23. the equity in application principle

24. the less worse alternative principle

25. the sometimes it's better to do nothing principle

Table 5. Questions for judgment and policy

1. *Goals*—have the goals been clearly stated, justified and prioritized? Are they consistent with the values of a democratic society?

2. *Accountable, public and participatory policy development*—has the decision to apply the technique been developed through an open process, and if appropriate, with participation of those to be surveilled? This involves a transparency principle.

3. *Law and ethics*—are the means and ends not only legal, but also ethical?

4. *Opening doors*—has adequate thought been given to precedent-creation and long term consequences?

5. *Golden rule*—would the watcher be comfortable in being the subject rather than the agent of surveillance if the situation was reversed? If appropriate is their equity/equivalence in the relationship?

6. *Informed consent*—are participants fully appraised of the system's presence and the conditions under which it operates? Is consent genuine (i.e., beyond deception or unreasonable seduction) and can "participation" be refused without dire consequences for the person?

7. *Truth in use*—where personal and private information is involved does a principle of "unitary usage" apply in which information collected for one purpose is not used for another? Are the announced goals the real goals?

8. *Means-ends relationships*—are the means clearly related to the end sought and proportional in costs and benefits to the goals?

9. *Can science save us*?—can a strong empirical and logical case be made that a means will in fact have the broad positive consequences its advocates claim? (the "does it really work question")

10. *Competent application*—even if in theory it works, does the system (or operative) using it apply it as intended?

11. *Human review*—are automated results with significant implications for life chances subject to human review before action is taken?

12. *Minimization*—if risks and harm are associated with the tactic, is it applied to minimize these showing only the degree of intrusiveness and invasiveness that is absolutely necessary?

13. *Alternatives*—are alternative solutions available that would meet the same ends with lesser costs and greater benefits (using a variety of measures not just financial)?

Table 5. Questions for judgement and policy (continued)

14. *Inaction as action*—has consideration been given to the "sometimes it it is better to do nothing" principle?

15. *Periodic review*—Are there regular efforts to test the system's vulnerability, effectiveness, and fairness and to review policies?

16. *Discovery and rectification of mistakes, errors, and abuses*—are there clear means for identifying and fixing these (and in the case of abuse, applying sanctions)?

17. *Right of inspection*—can individuals see and challenge their own records?

18. *Reversibility*—if evidence suggests that the costs outweigh the benefits how easily can the surveillance be stopped (*e.g.*, extent of capital expenditures and available alternatives)

19. *Unintended consequences*—has adequate consideration been given to undesirable consequences, including possible harm to watchers, the watched, and third parties? Can harm be easily discovered and compensated for?

20. *Data protection and security*—can surveillants protect the information they collect? Do they follow standard data protection and information rights as expressed in the Code of Fair Information Protection Practices and the expanded European Data Protection Directive?