



## Data Quality, Database Fragmentation and Information Privacy\*

Martin R. Gibbs<sup>1</sup>, Graeme Shanks<sup>2</sup> and Reeva Lederman<sup>3</sup>

---

### Abstract

In this paper we use an Information Systems (informatics) perspective to critically examine legislation designed to regulate the way private sector organizations collect, store, use, and disclose personal information. We focus on The Privacy Amendment (Private Sector) Act 2000 (Cth), which has recently been enacted in Australia. We argue that the ability of organizations to respond to the requirements of this legislation is affected by the data quality of the personal information they possess. In particular, this paper examines one problem associated with data quality that erodes an organization's ability to comply with legislation designed to protect the information privacy of individuals – the fragmentation of customer data across multiple databases owned and maintained by separate functional units within an organization. Given the ubiquity of these kinds of data quality problems we conclude that current legislative regimes to regulate private sector use of personal information in countries such as Australia and European Union member states can lead to contrary outcomes resulting in legislation that is either unenforceable or acts to encourage the development of high-quality, integrated customer databases that have the potential to erode information privacy. We believe that new models able to grapple with management of personal information in distributed, mobile and ubiquitous computing environments need to be developed.

---

### Introduction

The OECD's 1980 *Guidelines on the Protection of Privacy and the Transborder Flows of Personal Information* ([OECD 1980](#)) have provided a set of principles used by many countries to guide the formulation of information privacy regulations. These principles are intended to govern the ways in which organizations collect, store, use and disclose personal information. In this paper we focus on legislation derived from these principles that has recently been enacted in

---

\* An earlier version of this paper was presented at the Third Australian Institute of Computer Ethics Conference (AiCE'02), Sydney, School of Information Technology, 30 Sept. 2002. Thanks to Roselle de Silva for her assistance in this project, Sean Smith for encouragement and thoughtful comments, the anonymous reviewers for their critical insights, and to the participants at AiCE'02 for their feedback on an earlier version of this paper.

<sup>1</sup> Department of Information Systems, University of Melbourne, Australia. <mailto:martinrg@unimelb.edu.au>

<sup>2</sup> School Of Business Systems, Monash University, Australia.

<sup>3</sup> Department of Information Systems, University of Melbourne, Australia.

Australia: the Privacy Amendment (Private Sector) Act 2000 (Cth), which that came into effect on the 21<sup>st</sup> of December 2001. The Private Sector Amendment applies to the vast majority of private sector organizations with annual turnovers of \$3 million or more and to organizations that provide health services or hold health related information. It also applies to organizations with smaller annual turnovers that trade in personal information. The Private Sector Amendment regulates the ways in which these private sector organizations can collect, store, use and disclose personal information and it gives individuals the legal right to access and correct information held about them by private sector organizations ([OFPC 2001a](#)). As such, the Private Sector Amendment presents a number of challenges to organizations that collect, use and distribute personal information. In order to meet these challenges, many organizations will have to change the way they handle personal information.

It is now commonplace for commercial organizations to collect information about their customers and to compile extensive databases containing personal details as well as information such as consumer preferences, purchasing habits, medical conditions and so forth ([Lyon 1994](#)). These customer databases are important repositories of personal information for many organizations. Previous studies have shown that maintaining consistently high levels of customer data quality in these customer databases is a significant challenge and considerable expense for organizations ([Redman 1998](#), [Strong et al. 1997](#), [Wang 1998](#) and [Etzioni 1999, 134-6](#)). We argue that the ability of organizations to comply with the provisions of the Private Sector Amendment will be significantly compromised by the data quality of the personal information they hold. This connection between poor customer data quality and information privacy is clearly important in determining an organization's ability to comply with legislation designed to protect information privacy ([Gibbs et al. 2002](#), [Lederman, Shanks and Gibbs 2002](#)).

In this paper we present findings from an exploratory study designed to identify significant issues created by poor customer data quality that face organizations as they adjust their business practices to meet the provisions of new privacy legislation. The next section of this paper briefly describes the provisions of the Private Sector Amendment. Following this description we outline the connection between information privacy and the data quality of personal information. We then briefly describe our research approach before moving on to present and discuss some significant results from our study. In concluding this paper we observe that poor customer data quality erodes an organization's ability to *control* the personal information it possesses. This erosion of control can seriously hamper an organization's ability to comply with the provisions of the new legislation. Although the explicitly stated intention of the Private Sector Amendment is to 'give people some control over the way information about them is handled' ([OFPC 2001b](#)), this control cannot be secured by individuals unless organization's have control of this personal information in the first place. Thus, poor data quality has significant implications for the protection of information privacy that extend beyond ensuring that personal information is complete, accurate and up-to-date.

While this paper focuses on Australian privacy legislation, our findings have a wider significance. The Privacy Act and its later amendments are based on the OECD's 1980 *Guidelines on the Protection of Privacy and the Transborder Flows of Personal Information* ([OECD 1980](#)). Therefore, the issues we identify are not limited to the Australian context but can be generalised

to all organizations that must comply with privacy laws derived from these OECD principles such as those currently being enacted by member nations of the European Union in response to the 1995 European Union (EU) Data Protection Directive (EU 1995).

### **The Privacy Amendment (Private Sector) Act 2000**

Most definitions of privacy invoke one or more of the following three key elements: anonymity, solitude and/or secrecy (See for example [Johnson 2001](#) and [Spinello 2000](#)). These elements are often expressed as rights of the individual: the right to act anonymously, the right to live free of unwanted harassment, and/or the right for individuals to choose how they present themselves to others. When discussing privacy and information technologies, the last of these listed rights is often restated as the right for individuals to control the access others have to their personal information. Given that a number of possible definitions of privacy have common currency it is interesting to note that neither the Australian Commonwealth Privacy Act 1988 (Cth) (Privacy Act) nor its more recent Private Sector Amendment explicitly define privacy.

While privacy is not explicitly defined within these pieces of legislation, they do provide sets of 'privacy principles' for the protection of personal information. The Privacy Act is based on OECD principles for the protection of personal information. Prior to 2001, it regulated the handling of personal information by Commonwealth Government departments and credit reporting organizations. It also established the office of the Federal Privacy Commissioner. The Private Sector Amendment to this act extended privacy protection legislation to the private sector.

The Federal Privacy Commissioner has explicitly stated on numerous occasions that the goal of the Private Sector Amendment is to give individuals 'some control' over the personal information held about them by private sector organizations ([FPC 2000](#), [OFPC 2001a](#), [2001b](#)). As such, it is useful to understand the definition of privacy implicit within the Privacy Act as pertaining solely to 'information privacy'. Roger Clarke ([1999](#)) has usefully defined information privacy in the following way:

Information privacy refers to the claims of individuals that data about themselves should generally not be available to other individuals and organizations, and that, where data is possessed by another party, the individual must be able to exercise a substantial degree of control over that data and its use.

The Privacy Act regulates the way personal information is collected, stored, used and disclosed. Personal information is defined within the Privacy Act as:

Information or an opinion (including information or an opinion forming part of a database), whether true or not, and whether recorded in a material form or not, about an individual whose identity is apparent, or can reasonably be ascertained, from the information or opinion. (The Privacy Act 1988 (Cth), Sect 6)

The Private Sector Amendments established ten National Privacy Principles (NPPs) as the minimum standard for information privacy in the private sector ([AGD 2001](#)). The NPPs govern how an organization should handle personal information. They cover: collection (NPP1); use and disclosure (NPP2); data quality (NPP3); data security (NPP4); openness (NPP5); access and correction (NPP6); use of government identifiers (NPP7); anonymity (NPP8); transborder data flows (NPP9); and sensitive information (NPP10). The NPPs encapsulate a similar set of principles to those encapsulated in the OECD's 1980 principles for information privacy protection and the 1995 EU Data Protection Directives.

The Private Sector Amendment is one of several recent measures introduced by the Australian Government to facilitate Australia's transition to an information economy ([DCITA 2000](#)). The new provisions have been implemented with the stated aim of balancing individual's rights for information privacy against the 'right of government and business to achieve their objectives in an efficient way' ([FPC 2000, 2](#)). In the spirit of promoting a 'culture that respects privacy' ([FPC 2000, 2](#)), privacy has been promoted as being 'good business' ([OFPC 2002](#)) as well as being good for individuals. In particular, the legislation has been implemented with the recognition that consumers' lack of trust in the way commercial organizations handle their personal information is a major barrier to the growth of e-commerce ([OFPC 2002](#)). The legislation also represents an attempt to bring Australia into line with international privacy regimes especially those of the European Union (EU) given the possibility that the EU will impose trade restrictions on nations that do not adequately protect the personal information of EU citizens.

In developing the provisions of the Private Sector Amendment the Australian government has deliberately opted for a 'light-touch' co-regulatory approach to the regulation of privacy with the aim of encouraging compliance through facilitation rather than through the threat of punitive actions for non-compliance ([OFPC 2001c](#), [OFPC 2001d](#)). This approach has been designed to minimise the burden of compliance for businesses. It is also an approach that has attracted strong criticism and has led to the amendments being dubbed 'anti-privacy laws' (Roger Clarke quoted in [Haslam and Mitchell 2001](#)) and described as 'reducing existing privacy protection' ([Clarke 2000](#)) due to the large number of exceptions and qualifications built into the legislation and because it seemingly 'legitimises many unreasonable uses of personal data' ([Clarke 2000](#)). The legislation has also been criticised for lacking 'grunt' and being 'toothless' due to the Federal Privacy Commissioner not being granted significant investigative powers or an ability to impose significant punitive penalties for breaches of the Privacy Act ([McClelland in Australia, House of Representatives 2000, 22233-7](#)). While these criticisms are significant and have some bearing on how private sector organizations have responded to the new privacy provisions, they are not the major focus of this paper. Rather, we wish to approach the question of information privacy from an Information Systems perspective that critically examines information privacy protection legislation in terms of database 'informatics'. Specifically, we wish to examine how poor data quality in personal information will affect an organization's ability to comply with the provisions set out in the Privacy Act and its recent amendments.

## Data Quality

NPP3 of the Private Sector Amendment sets out expectations for the maintenance of data quality. It requires an organization to ensure that the personal information it collects, uses or discloses is accurate, complete and up-to-date. This approach to data quality is quite typical. Much of the existing work on data quality focuses on the intrinsic quality of data in databases and consists of lists of desirable information quality dimensions ([Wand and Wang 1996](#)). These lists typically include dimensions such as accuracy, completeness as well as reliability, consistency, timeliness, precision and conciseness ([Wang and Strong 1996](#), [Kahn et al. 2002](#)).

As such, these data quality frameworks focus primarily on the content or ‘meaning’ of particular data fields. That is, they define data quality in terms of the data’s *semantic* properties. However, a wider or more rounded view of data quality can be adopted: a view of data quality that defines quality in terms of the data’s ‘fitness’ for particular purposes or organisational functions ([Shanks and Darke 1998](#)). Viewed in this manner, data quality can be seen to involve more than semantic accuracy and completeness. Data has other characteristics such as its structural properties, its useability, and its openness to multiple interpretations in different contexts that also determine its quality or fitness for particular purposes ([Shanks and Darke 1998](#)).

Price and Shanks ([2004](#)) have developed an analytic framework based on semiotic theory for the study of data quality in this manner. Their semiotic framework has three discrete levels of analysis: syntactic, semantic and pragmatic. Within this framework *syntactic* data quality refers to the data structures used to store personal information. Syntactic data quality is a measure of the consistency of representation in one or more databases. *Semantic* data quality focuses on the meaning of data and measures how complete, accurate and up-to-date it is. *Pragmatic* data quality is concerned with the utility of data for specific tasks and is a measure of the usefulness and useability of data. It will vary with the person involved, the task at hand and the organisational context of use.

This semiotic framework provides a set of generative concepts and analytic distinctions that are useful for investigating and understanding some of the impediments and problems face by organization as they move to comply with the provisions of new privacy legislation. It is particularly useful because, by defining data quality in terms of fitness for purposes, it focuses our attention on the connection between the qualities of an organization’s databases and the organization’s ability to respond to the regulatory requirements of privacy legislation.

## Research Approach

This research study was exploratory in nature and involved two main phases: a conceptual phase and an empirical study. The conceptual study phase of the research included a critical review of the Privacy Act and Private Sector Amendment, associated submissions to parliament, press commentary and other relevant literature from both academic and practitioner sources. This material was then synthesised with concepts from Price and Shanks’ ([2004](#)) semiotic framework for understanding data quality in order to develop an initial understanding of how poor customer

data quality may prevent organizations from fulfilling their obligations to maintain the information privacy of individuals, and to develop an interview protocol for data collection in the empirical phase of the research.

The empirical phase of the study involved in-depth interviews with eight experienced practitioners. Interviewees were identified opportunistically and selection for interview was based on the criteria that they had extensive experience with information privacy and the management of information systems. Five of our interviewees occupied senior, information systems management roles in private sector organizations that handled large amounts of customer data. The other three were consultants specialising in the areas of privacy and/or data management. Empirical data was collected through open-ended and semi-structured interviews and review of documents contributed by interviewees. Interview duration ranged from 60 to 90 minutes and were recorded on audiotape and fully transcribed. Transcripts were used to identify key issues associated with data quality faced by organizations as they responded to the provisions of the legislation. From this list of key issues that have previously been discussed elsewhere ([Gibbs et al. 2002](#)) one issue has been selected for further elaboration in this paper on the basis of its relevance, importance and frequent occurrence. In this paper we discuss the problems created by the fragmentation of personal information about an individual across a number of different databases that are maintained and controlled by different function units within an organization. In particular, we will discuss the difficulties this situation creates for organizations when called upon by customers to complying with the provisions of NPP 6 (access and correction) and NPP 3 (data quality).

## Data Fragmentation and the Control of Personal Information

Many large organizations of the type covered by the Private Sector Amendment have a history of separate business units developing and maintaining independent customer databases. Typically these legacy systems will have been developed autonomously and use a variety of data structures and identifiers to record personal information. In addition, these databases are often ‘owned and operated’ by separate functional units within the organization. Consequently, the personal information an organization holds about individuals will be fragmented across a number of databases using a variety of different data structures ([Redman 2001](#)). This makes accessing and collating personal information difficult and time-consuming ([Redman 1998](#)). Rarely in these cases is there a unified and consolidated view of the personal information an organization holds about an individual ([Shanks 1997](#)).

These kinds of data quality issues that extend across the syntactic level (in the form of incompatible and inconsistent data structures) to the pragmatic level (in the form of data that has low useability and usefulness) make it difficult for an organization to comply with some of the provisions of the Private Sector Amendment. In particular and most strikingly, it creates problems with the central information privacy tenet associated with giving individuals control over their personal information: allowing individuals to access and correct personal information held about them by an organization. This principle is codified in NPP 6 of the Private Sector Amendment. NPP 6 – Access and Correction – states that an organization must give individuals

access to their personal information if requested and they must correct that information if it is inaccurate, incomplete or out-of-date.

One of our interviewees, the information systems manager for a major metropolitan hospital, reported that locating and identifying all the databases within the organization that contained identifiable personal information was a major problem for his organization's ability to comply with the new privacy legislation. While the paper-based patient record recorded all treatment that patients received within the hospital, various units within the hospital also maintained their own, separate records for a variety of purposes associated with research, treatment and service evaluation as well as for the purposes of providing health services to the patient. In addition, some senior specialists who consulted with patients in the hospital maintained their own, private records and notes on patients independent of the main hospital's patient record system. Senior specialists and the central medical record and patient billing systems aside, this organization had approximately 30 different function units that collected and used personal information; many of them using more than one information system to do so. While a portion of these information systems were modest in scale – spreadsheet applications and small databases on desktop workstations and personal computers – the difficulties this situation created for the organization in compiling a view of the totality of personal information held about any one individual, are obvious. This degree of fragmentation creates serious pragmatic data quality problems and had severe implications for this organization's ability to respond in a timely and efficient manner to an individual's request to access their personal information as required by NPP6 (Access and Correction). This fragmentation also generates enormous difficulties for ensuring that all personal information held by this organization was accurate, complete and up-to-date at the semantic data quality level as required by NPP3 (Data Quality).

Three other interviewees who worked for two large retail organizations also reported problems of a similar nature. Although the problems were on a smaller scale, the organizations they worked for were grappling with similar issues associated with multiple and fragmented databases. In both cases, despite having made significant moves towards consolidating customer databases used for a variety of purposes such as tracking purchasing habits, marketing, lay-buy, in-store credit facilities and valued customer schemes, these organizations still had personal information about customers contained in multiple databases owned and maintained by different departments within each of these organizations.

One of these organizations, a large high street fashion retail chain, maintained a number of databases related to its credit schemes, loyalty programs and other marketing activities. Personal information was collected in a variety of ways for these purposes, often at the point-of-sale. Some of this information for promotional and marketing activities was maintained locally, at the retail outlet and existed outside of any system of centralized control, while others were coordinated centrally. In addition, personal information associated with lay-buys (lay-away purchases) was typically stored in small database systems operated by the shop-front staff and maintained locally at the retail outlet.

The other organization, a large department store, maintained an integrated, centrally controlled customer relationship management (CRM) system. Yet, it too had problems with database

fragmentation as several departments within the organization such as human resources, marketing and in-store security insisted on maintaining their own databases that contained significant amounts of personal information. These departments also insisted that other departments within the organization kept their 'hands-off' these information systems and thus these databases that contained personal information were not subject to any form of unified or centrally coordinated control by the organization. For example, the marketing department maintained separate lists of personal details for a number of marketing activities separate from the centrally coordinated CRM system. These lists, which contained significant personal information, were maintained on desktop computers in the marketing department. They were used to 'wash' data from the CRM in order to generate new lists for activities such as direct marketing and offers of promotional opportunities to customers. The in-store security department also maintained its own databases of 'known' and suspected shoplifters. This information was regularly exchanged with the security departments of other major retail outlets in the area. The fragmentation and scattering of personal information across this organization and the lack of control over personal information it created had particularly serious implications in regard to the sensitive nature of the personal information and opinions about individuals being collected, stored, used and disclosed to other organizations by the in-store security department.

Although no customers of these two retail organizations had requested access and correction of the personal information held by these organizations at the time of interview, our interviewees anticipated that this lack of cohesion in their databases would create significant difficulties with providing customers with access to all the personal information held about them and would impede their organization's ability to comply with the access and correction provisions of the Privacy Act. They also expected to encounter significant difficulties ensuring that all personal information held by their organization was semantically accurate, complete and up-to-date.

The spread of these kinds of problems across the private sector were confirmed by the three consultants we interviewed based on their experience with a broad range of private sector organizations that operated in a variety of different sectors of the economy.

## Discussion

A review of Australian privacy legislation indicates that this body of legislation is based on the assumption that organizations have an integrated customer data set and that it is relatively easy for organizations to access, collect and collate all the personal information they hold about an individual. However, the situation is quite different for many organizations. Often organizations do not have integrated customer data sets ([Redman 1998](#), [Shanks 1997](#), [Strong et al. 1997](#), [Wang 1998](#)). Thus, we would expect that these organizations cannot readily achieve the unified view necessary for strict and unproblematic compliance with the provisions of this legislation due to problems with data quality. The results of our research suggest that this expectation is indeed the case for many organizations.

This kind of data quality problem is particularly pernicious for organizations with multiple points of customer contact. These organizations are often characterised by semi-autonomous functional

units that have been in the habit of amassing their own customer databases without reference to a centrally coordinated information management strategy. As a result, the sum total of personal information held about any individual is fragmented across multiple and incompatible databases creating significant data quality problems that severely hinder the formation of a unified and integrated view of the personal information held about any one particular individual. This inability to develop an integrated view compromises an organization's ability to effectively manage and control its customer data and hence compromises its ability to meet its obligations under the Privacy Act and its recent amendment.

The ability to develop a unified and integrated view of the totality of personal information held about an individual should enable organizations to comply with the provisions of the new privacy legislation in a relatively unproblematic fashion. Establishing and maintaining high levels of customer data quality across all three data quality levels is an important part of developing this kind of view of the personal information held about a particular individual. It is ironic to note that it is precisely these kinds of information systems that are built on good quality, highly integrated databases of personal information that have raised the hackles, suspicions and fears of surveillance studies academics, privacy advocates and political commentators for several decades due to their ability to enable organizations to practice data-mining and data-matching activities (See for example [Clarke 1988](#), [Davies 1997](#), [Regan 1995](#)). Yet, it would seem, it is precisely those organizations with highly integrated and carefully managed customer databases that are in the best position to comply with the provisions of current information privacy legislation.

Given current legislative regimes to regulate private sector use of personal information in countries such as Australia and European Union member states, the *legal* protection of information privacy may well be better served by tightly controlled, monolithic, centrally coordinated databases than less well served by these forms of data storage. This, we would argue, is due to the relative ease of compliance with core information privacy protection principles such as accuracy, access and correction enabled by these kinds of integrated and coherent databases. However, it is with good reason that privacy advocates and surveillance studies academics have cautioned against these kinds of databases. These kinds of databases pose a significant threat to information privacy because of the organizational practices they enable. High quality integrated databases are an important organizational resource that enable a host of valuable organizational practices that have been shown to reduce operating costs, improve customer service, and provide decision support ([Redman 1998](#), [Strong et al. 1997](#) and [Wang 1998](#)). Not inconsequentially, they are also a necessary resource for performing information privacy eroding activities such as data merging, matching and mining practices. As a result, legislation designed to protect information privacy may produce two contrary outcomes.

First, legislation designed to protect information privacy may encourage organizational practices that lead to an erosion of information privacy. That is, privacy legislation that requires organizations to allow individuals to access and correct personal information could force organizations to address their database fragmentation problems. Successfully addressing these data quality problems should produce integrated and coherent data repositories of the kind that pose to the most threat to information privacy if misused. This is particularly true of legislation

that is overly punitive. If the costs of non-compliance to organizations are high, then strict adherence to the letter, rather than the spirit, of the law may be promoted. It is worth noting that in developing the Private Sector Amendment the Australian government deliberately opted for a 'light-touch' co-regulatory approach to the regulation of information privacy. The stated aim of this approach was to protect privacy by encouraging and promoting a 'culture that respects privacy' ([FPC 2000, 2](#)) rather than seeking to enforce privacy protection through the threat of punitive actions for non-compliance ([OFPC 2001c](#), [OFPC 2001d](#)). Thus, despite its critics, it is possible to speculate that the 'soft-touch' approach adopted in Australia may be more appropriate, and lead to better protection of information privacy, than would a heavy-handed, punitive approach.

Secondly, and contrary to the intent of information privacy legislation, there is a danger that legislation of this kind is used by organizations to justify the very organizational practices it was designed to inhibit. That is, the need to comply with legal requirements for access and correction of personal information could be used to justify the inappropriate consolidation and centralized coordination of personal information from multiple business units that are all part of the same corporate entity. It is not difficult to imagine a large corporation seeking to combine and consolidate personal information from business units and subsidiaries engaged in a range of separate business activities. Rather than creating the opportunity to justify the consolidation of personal information, information privacy legislation needs to ensure that privacy bulkheads limiting the exchange of personal information between operating units in large multi-business organizations remain in place.

However, many organizations such as those we examined in our study are large and have multiple points of contact with their customers, yet they can still be regarded as engaging with customers in only one arena of their lives. In these circumstances it is not unreasonable to assume that a request for access to personal information should encompass all the personal information held by the organization. For example, when approaching a hospital, it is reasonable to expect that all personal information collected by the hospital's various clinics and consultants would be available along with the centralized patient record. Similarly, retail organizations should be able to provide a consolidated view of the personal information held about an individual that encompasses the marketing, finance, and security departments as well as personal information that might be held at local retail outlets. In these circumstances, it is unreasonable to expect an individual to approach all the separate units of the organization to access the personal information held by the organization. This, we believe, suggests that a more nuanced articulation of these privacy principles in law is needed, such that the legislation cannot be used to justify the inappropriate consolidation of personal information from multiple business units engaged in diverse areas of operation, while at the same time, still ensuring that all personal information held about an individual by a single organizational 'entity' can be unproblematically accessed and corrected if requested. The determination of what constitutes a single organizational entity for these purposes should be based on the expectations of a reasonable person in the context of the request and limited to units engaged in closely related activities. These problems of aggregation and the lack of clearly articulated guidelines to protect against inappropriate consolidation of personal information across multiple agencies are not limited to Australian privacy legislation but have also been highlighted in critiques of the OECD principle themselves ([Clark 2000b, S4.3](#)).

## Conclusion

It is apparent that although the explicitly stated intention of the Private Sector Amendment is to ‘give people some control over the way information about them is handled’ (OFPC 2001b), poor customer data quality in the form of database fragmentation can undermine the ability of an organization to manage and use, that is control, the personal information it holds. In order to cede or ‘give control’, one must have control in the first place. Data quality is an important factor in determining the amount of control organizations have over the personal information they hold. Poor data quality erodes an organization’s ability to control the personal information it holds about individuals and this erosion inhibits their ability to comply with the requirements of this privacy legislation. Establishing and maintaining high levels of customer data quality is therefore a necessary and potentially expensive step that any organization will need to take in order to be able to fully comply with current information privacy laws. This necessity extends beyond maintaining data quality at the semantic level by ensuring that personal information is complete, accurate and up-to-date as required by NPP3 – data quality. Data quality must also be maintained at the syntactic and pragmatic levels if an organization is to have sufficient control over the personal information it holds to be able to give individuals the ability to access and correct this information. Of course, in addition to enabling access and correction of personal information, data quality is also an important prerequisite for the performance of a wide variety of organization activities and functions including activities that pose a threat to information privacy such as data merging, matching and mining. Data quality is thus a two-edged sword as far as privacy legislation based on OECD data protection principle is concerned. Good quality databases are needed for organizations to be able to comply with these principles, yet they also enable organizational practices that have the potential to erode information privacy if misused.

Our study suggests that privacy legislation based on OECD principles formulated in the 1970s and early 1980s does not adequately grapple with the reality of technological development that has occurred over the last two decades. In particular, privacy legislation premised on the assumption that personal information is stored on databases that are centralized resources within an organization no longer holds. The spread of desktop personal computers in the 1980s and of networking technologies in the 1990s has distributed and dispersed computing resources throughout many organizations, and with it, personal information. The rapid and recent spread of mobile and ubiquitous computing devices will only exacerbate the problems of fragmentation and scattering of personal information we have identified, making compliance even more difficult. Perhaps, as David Lyon (2002) has recently suggested, metaphors and models such as Bentham’s Panopticon and centrally coordinated, tightly integrated monolithic databases may no longer be particularly useful for understanding surveillance and threats to privacy in contemporary society. New metaphors and models able to grapple with the decentring and distributed possibilities of these new technologies and able to inspire appropriate and effective legislative action are clearly required.

## References

- Attorney-General's Department (AGD) (2001) Extract from the Privacy Act 1988: The National Privacy Principles in the Privacy Amendment (Private Sector) Act 2000 as at 10/01/2001, 19 December 2003. <http://www.privacy.gov.au/publications/npps01.pdf>
- Australia, House of Representatives (2000) *Votes and Proceedings*: 22233-7.
- Clarke, R. (2000a) Privacy Bill Needs Much More Work, *The Australian*, 15 Feb 2000, 17 Dec 2003. <http://www.acs.org.au/news/oz150200.htm>
- Clarke, R. (2000b) Beyond the OECD Guidelines: Privacy Protection for the 21<sup>st</sup> Century. Xamax Consultancy, Canberra. <http://www.anu.edu.au/people/Roger.Clarke/DV/PP21C.html>
- Clarke, R. (1999) Internet Privacy Concerns Confirm the Case for Intervention, *Communications of the ACM*, 42(2): 60-67.
- Clarke, R. (1988) Information Technology and Dataveillance, *Communications of the ACM*, 31(5): 498-512.
- Davies, S. (1997) Re-Engineering the Right to Privacy: How Privacy Has Been Transformed from a Right to a Commodity in P. Agre and M. Rotenberg (eds.) *Technology and Privacy: The New Landscape*, MIT Press, Cambridge Mass.
- Department of Communications Information Technology and the Arts (DCITA) (2000) Submission from the Department of Communications Information Technology and the Art to the Senate Select Committee on Information Technologies Inquiry into e-Privacy, July 2000.
- EC/95 (1995) Directive 95/46/EC of the European Parliament On the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, 24 October, 1995.
- Etzioni, A. (1999) *The Limits of Privacy*. New York: Basic Books.
- Federal Privacy Commissioner (FPC) (2000) Submission from the Federal Privacy Commissioner to the House of Representatives Standing Committee on Legal and Constitutional Affairs Inquiry Into the Privacy Amendment (Private Sector) Bill 2000.
- Gibbs, M.R, G. Shanks, R. Lederman and R. de Silva (2002) Privacy and Customer Data Quality: Exploring the Issues. In *Enabling Organisations Through Information Systems: Proceedings of the 13th Australasian Conference on Information Systems*, Melbourne, 4-6 December 2002: 279-292.
- Haslem, B. and Mitchell, S. (2001) New Laws 'Complex and Full of Holes' *Australian IT*, 21 December 2001, 2 Jan 2002. <http://austalianit.news.com.au/>
- Kahn, B., Strong, D.M. and Wang, R.Y. (2002) Information Quality Benchmarks: Product and Service Performance, *Communications of the ACM*, 45(4): 184-192.
- Johnson, D.G. (2001) *Computer Ethics*. Upper Saddle River NJ: Prentice Hall.
- Lederman, R., G. Shanks and M.R. Gibbs (2003). Meeting Privacy Obligations: The Implications For Information Systems Development. In *New Paradigms in Organizations, Markets and Society: Proceedings of the 11th European Conference on Information Systems*. Naples 19-21 June, 2003.
- Lyon, D. (2002) Surveillance Studies: Understanding Visibility, Mobility and the Phenetic Fix. *Surveillance & Society*, 1(1): 1-7. <http://www.surveillance-and-society.org/articles1/editorial.pdf>
- Lyon, D. (1994) *The Electronic Eye: The Rise of Surveillances Society*. Oxford: Polity Press.

- Organisation for Economic Co-operation and Development (OECD) 1980 *Guidelines on the Protection of Privacy and the Transborder Flows of Personal Information*. 23 September 1980, 18 December 2003. <http://www1.oecd.org/publications/e-book/9302011E.PDF>
- Office of the Federal Privacy Commissioner (OFPC) (2002) *Good Privacy, Good Business: Privacy in Australia*, AGPS, Canberra, November 2002, 17 December 2003. <http://www.privacy.gov.au/publications/pianew.pdf>
- Office of the Federal Privacy Commissioner (OFPC) (2001a) *Information Sheet 1-2001: Overview of the Private Sector Provisions*, December 2001, 17 November 2003, <[http://www.privacy.gov.au/publications/IS1\\_01.pdf](http://www.privacy.gov.au/publications/IS1_01.pdf)>
- Office of the Federal Privacy Commissioner (OFPC) (2001b) *Information Sheet 2-2001: Preparing for 21 December 2001*, December 2001, 17 November 2003. [http://www.privacy.gov.au/publications/IS2\\_01.pdf](http://www.privacy.gov.au/publications/IS2_01.pdf)
- Office of the Federal Privacy Commissioner (OFPC) (2001c) *Information Sheet 13-2001: The Privacy Commissioner's Approach to Promoting Compliance with the Privacy Act*, December 2001, 17 November 2003. [http://www.privacy.gov.au/publications/IS13\\_01.pdf](http://www.privacy.gov.au/publications/IS13_01.pdf)
- Office of the Federal Privacy Commissioner (OFPC) (2001d) *Implementation of the Privacy Amendment (Private Sector) Act 2000, Privacy Law and Policy Reporter 8*, online, 12 April 2002. <http://www.austlii.edu.au/au/journals/PLPR/2001/3.html>
- Price R.J., Shanks G. (2004) A semiotic information quality framework. In R. Meredith, G. Shanks, D. Arnott and S. Carlsson (eds.) *Proceedings of the 2004 IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World*, Prato, Italy, 1-3 July: 658-672.
- Redman, T. (2001) *Data Quality: The Field Guide*, New Jersey: Digital Press.
- Redman, T. (1998) The Impact of Poor Data Quality on the Typical Enterprise, *Communications of the ACM*, 41(2): 79-82.
- Regan, P.M. (1995) *Legislating Privacy: Technology, Social Values, and Public Policy*. Chapel Hill: The University of North Carolina Press.
- Shanks, G. (1997) The Challenges of Strategic Data Planning in Practice: An Interpretive Case Study, *Journal of Strategic Information Systems*, 6(1): 69-90.
- Shanks, G. and Darke, P. (1998) Understanding Data Quality in Data Warehousing: A Semiotic Approach in I. Chengilar-Smith and L. Pipino (eds.) *Proceedings of the International Conference on Information Quality*, MIT, Boston, November: 247-264.
- Sinclair, J. (2002) The Charge to See, *The Age*, 18 February, Money Manager Suppl.:3.
- Spinello, R.A. (2000) *Cyberethics : Morality and Law in Cyberspace*. Boston, MA: Jones and Bartlett
- Strong, D.M., Lee, Y.W. and Wang, R.Y. (1997) Data Quality in Context, *Communications of the ACM*, 40(5): 103-110.
- Wand, Y. and Wang, R. (1996) Anchoring Data Quality Dimensions in Ontological Foundations. *Communications of the ACM*, 39(11): 86-95.
- Wang, R.Y. (1998) A Product Perspective on Total Data Quality Management. *Communications of the ACM*, 41(2): 58-65.

Wang, R.Y. and Strong, D.M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers.  
*Journal of Management Information Systems*, 12(4): 5-34.