

Meeting Privacy Obligations: The Implications For Information Systems Development

Reeva Lederman

Department of Information Systems University of Melbourne, Victoria, 3010

Tel: +613 83441535, Fax: +613 9349 4596

r.lederman@dis.unimelb.edu.au

Graeme Shanks

Department of Information Systems University of Melbourne, Victoria, 3010

Tel: +613 83441577, Fax: +613 9349 4596

g.shanks@dis.unimelb.edu.au

Martin R Gibbs

Department of Information Systems University of Melbourne, Victoria, 3010

Tel: +613 83441394, Fax: +613 9349 4596

m.gibbs@dis.unimelb.edu.au

Abstract

The European Union (EU) Data Protection (Privacy) Directive of 1995 (EUPD) and resulting legislation introduced by member states is designed to ensure that business activity is subject to privacy regulation. The ability of organisations to respond to the requirements of this legislation is affected by the quality of their customer data. This paper explores the issues for IS development created by poor customer data quality as organisations adjust their business practices to meet the new legislative provisions. A number of key issues emerge including managing large amounts of fragmented customer data, understanding what information is required for organisational activities, controlling use and disclosure across the organisation, and allowing anonymity when interacting with customers. Furthermore, several important implications for systems development practitioners are discussed.

Keywords

Data Privacy, data quality

1. Introduction

Prior to 1995 the constitutions of many EU countries recognised the right of privacy and private communications. The Belgian Constitution states, for example, that 'Everyone has the right to the respect of his private and family life.' (Article 22) The Danish and Finish constitutions also had similar provisions. Many EU countries such as France, Germany and Great Britain have less specific provisions although a number have provided Common Law protection for privacy (eg. *Halford v United Kingdom* 1997).

The European Union Data Protection Directive (EUPD) introduced by the Union Council on February 20 1995 ((EC)/95) attempted to consolidate the fragmented protections available across the Union by imposing one uniform set of directives to which all member countries were obliged to adhere.

While European Union directives have an immediate direct effect on business organisations since they apply without member states having to pass their own legislation, in fact most member states have introduced their own privacy protection legislation between 1996-2000, based on the provisions of the EUPD. For example, the Belgian Data Protection Act gives private companies and government agencies more specific direction in the application of the EU provisions, as does the Federal Data Protection Law in Germany and the Data Protection Act of the French Republic.

These various Acts present a number of challenges to organisations that collect, use and distribute personal information. Meeting these challenges has required significant changes to the ways in which organisations handle personal information (Davies 1997) and has created new responsibilities in IS development for organisations to maintain the data quality of the personal information they hold.

The main repository of personal information in many organizations is customer databases. Previous studies have shown that maintaining consistently high levels of customer data quality is a significant challenge and considerable expense for organisations (Redman 1998, Wang 1998). In this paper we argue that the ability of organisations to comply with the provisions of the EUPD will be significantly influenced by the data quality of the personal information they hold. In particular, poor customer data quality will create a number of serious problems for organisations in this regard. The connection between poor customer data quality and privacy is one that has not been explored in any detail previously, yet is clearly of great importance to EU organisations seeking to comply with relevant legislation and the EUPD.

Our case studies were conducted in Australia where pressure to reform privacy legislation emerged to a significant extent as a response to the EU directive to enable trans-border flows of personal information between Australia and EU countries. This Australian legislation has enacted a set of privacy principles very similar to those of the EUPD. Thus, the issues we identify in this study can be generalised to organizations in countries that must comply with privacy laws derived from these EU principles.

This paper first outlines the provisions of the EUPD. The following section describes a semiotic framework for understanding data quality (Shanks and Darke 1998), and relates relevant privacy principles to data quality dimensions. A discussion of the research approach used in this exploratory study follows. The next section includes a detailed discussion of the key issues that emerged from the study concerning how poor customer data quality impacts the ability of organisations to comply with the provisions of the EUPD. We observe that although the intention of the EUPD is to give people some control over the way information about them is handled, poor customer data quality degrades an organisation's ability to effectively manage and use the information it holds. This degradation in the

control over the personal information an organisation possesses, undermines its ability to cede some of that control back to the individuals concerned with serious implications for the protection of their information privacy. A number of important implications for information systems' practice conclude the paper.

2. Background

The EUPD established a number of privacy principles as the minimum standard for information privacy in the private sector and the legislation of member states such as France and Germany mentioned earlier further encode these principles.

Under the heading 'General Rules for the Lawfulness of the Processing of Personal Data' we find the general principles governing how an organisation should handle personal information.

In relation to the current study concerning the impact of poor customer data quality, six of these categories of principles are relevant and are paraphrased below:

Article 6.1.b – Collection: Data must be collected only for specified, explicit and legitimate purposes.

Article 6.1.c. – Use: Use of data must be adequate, relevant and not excessive.

Article 6.1.d – Data quality: Data must be accurate, up to date, complete, and able to be erased or rectified.

Article 6.1.e – Anonymity: Individuals should not be identified for longer than is necessary for the purposes for which the data is collected or further related processing.

Article 12.1 – Access: Individuals must be given access at reasonable intervals and without excessive delay or expense to personal data held by an organisation.

Article 12.2 – Correction: Incomplete or inaccurate data must be corrected through rectification, erasure or blocking of data.

Article 17 – Disclosure: Data must be protected against accidental or unlawful destruction or accidental loss and against unauthorised access, alteration, or disclosure.

In this paper we wish to discuss the unexplored issues associated with how poor data quality in personal information, with respect to the above principles, will affect an organisation's ability to comply with the provisions set out in the EUPD.

2.1 Customer Data Quality Management

Customer information is increasingly viewed by organizations as an important asset that can be used to deliver competitive advantage and support business initiatives that focus on the customer. Accordingly, it is crucial that the collection, storage and use of customer data is properly managed within organizations. A key aspect of customer data management is to ensure that the data is of high quality (Shanks and Tay 2001).

We define quality as 'fitness for purpose'. 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 completeness, accuracy, reliability, consistency, timeliness, precision and conciseness. Several frameworks have been developed that organise and structure important concepts in information quality (see for example Wand and Wang 1996, Kahn et al. 2002). In this paper, we use the framework of Shanks and

Darke (1998). This framework is soundly based in semiotic theory and includes both product-oriented and service-oriented aspects of data quality. *Semiotics* is the study of the use of symbols to convey knowledge and suggests four discrete levels of data quality: syntactic, semantic, pragmatic and social (Stamper 1992, Lindland et al. 1994).

Syntactic data quality is concerned with the structure of data. The goal of syntactic data quality is consistency of representation in one or more databases. For customer data, this includes consistent representation of such things as names and addresses and consistent use of coding schemes.

Semantic data quality is concerned with the meaning of customer data as assigned by users of the information. The goals of semantic data quality are that data is complete and accurate and up-to-date.

Pragmatic data quality is concerned with the use of customer data, and varies with the person involved, the task at hand and the organisational context. The goals of pragmatic data quality are usefulness and usability.

Social data quality concerns the shared understanding of data by various social groups within organisations or societies (Shanks and Corbitt 1999). It is relevant for organisations that have multiple points of contact with customers. The goals of social data quality are shared understanding of meaning and awareness of bias among different users. Using customer data without a shared understanding of its meaning and awareness of bias may lead to problems in correctly interpreting reports based on the data and problems with combining data from multiple sources.

Managing customer data quality involves understanding and measuring data quality problems and designing improvement strategies for both existing data stocks and incoming data flows (English 1999). Managers are increasingly asking for clear business benefits to be realised from expenditure on fixing data quality problems. Improving customer data quality leads to greater customer satisfaction (correct names and addresses, accurate billing, filled orders and receipt of appropriate marketing materials), decreased operational costs (less time and other resources spent detecting and correcting errors), more effective decision-making (accessing and using relevant and accurate information) and increased employee satisfaction (greater trust of information in databases) (English 1999, Redman 2001, Wand and Wang 1996). Increased data quality should also lead to greater ability to comply with the EUPD.

2.2 Research Question

The EUPD sets out expectations for the maintenance of data quality and requires organisations to ensure that the personal information it collects, uses or discloses is accurate, complete and up-to-date. That is, data quality is defined solely in terms of its semantic properties. However, if we adopt a broader view that understands quality as ‘fitness for purpose’, it is possible to see that data quality extends beyond the semantic dimension and has implications for many of the other requirements set out in the new legislation. That is, each relevant Article may be related to one or more data quality dimensions. Article 6.1.b (collection) is related to completeness at the semantic data quality level. Poor completeness indicates either missing data or excess data that is not required. This will violate 6.1.b. Articles 6.1.c and 17 (use and disclosure) constrain an organisation to only use or disclose personal information for the primary purposes it was collected unless the individual concerned has given their consent. This is related to usefulness at the pragmatic data quality level. Data that is useful will support the activities it was intended to support. Article 6.1.d (data quality) ensures that the personal information an organisation collects, uses or discloses is accurate, complete and up-to-date. This is the complete set of data quality dimensions at the semantic data quality level. Articles 12.1

and 12.2 (access and correction) concerns giving individuals access to their personal information if requested and allowing them to correct it. This is related to the accessibility data quality dimension at the pragmatic data quality level. Article 6.1.e (anonymity) concerns giving individuals the option to remain anonymous during any contact or transaction with the organization. This is related to the syntactic data quality level and concerns the representation of codes and identifiers.

Clearly, data quality issues are strongly related to the principles found in the EUPD. In this paper we will explore that relationship by identifying the data quality issues that occur in practice as organizations come to grips with their privacy obligations. That is, the research question addressed in this paper is:

What difficulties associated with data quality are organizations experiencing in their attempts to fulfil their obligations to maintain the information privacy of individuals as set out in the EUPD and what bearings might these difficulties have on IS development?

3. Research Approach

This research study involved a conceptual study and then an empirical study. The conceptual study phase of the research included an extensive review and synthesis of the EUPD, associated European legislation and comparison to the Australian legislation, press commentary and other relevant literature from both academic and practitioner sources. This was then synthesised with concepts from the 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 information privacy obligations, 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 privacy and data quality issues. Five interviewees had data management roles in organizations that handled large amounts of customer data in different industry sectors. The other three were consultants specializing in the areas of privacy and/or data management. Empirical data was conducted through semi-structured interviews and review of documents contributed by interviewees. Interview duration ranged from 60 to 90 minutes. Qualitative data analysis techniques (Miles and Huberman 1994) were used to identify key issues associated with data quality faced by organisations as they responded to the provisions of the EUPD-based Australian legislation. Given the similarities between the Australian-based legislation and that formulated on EUPD principles in European countries, we are confident the findings are relevant and applicable to the European environment.

4. Key Issues Emerging From The Study

Results of the data analysis suggested that the following four key issues were the most significant with regard to fulfilling privacy obligations:

4.1 Managing and Permitting Access to Large Amounts of Fragmented Customer Data

Many organisations have a history of separate business units developing and maintaining independent customer databases. Typically these legacy systems have been developed autonomously and use

different 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 organisation holds about individuals is fragmented across multiple and heterogenous databases. This makes accessing and collating personal information difficult and time-consuming (Shanks and Tay 2001, Strong et al. 1997).

One of our interviewees, the information systems manager for a large metropolitan teaching hospital, said that locating and identifying all the databases within the organisation that contained identifiable personal information was a major problem for her organisation'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 a health service to the patient. While a portion of these information systems were modest in scale – spreadsheet applications and small databases – the difficulties faced by the organisation as a whole 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 in the organisation's ability to respond in a timely and efficient manner to an individual's request to access their personal information as required by Articles 12.a and 12.b (Access and Correction).

4.2 Understanding What Personal Information is Required for Organisational Activities

Under Article 6.1.b (Collection), organisations may only collect personal information if it is required for a specific function or activity. Collecting personal information 'just in case' it is needed at some future point in time is no longer permissible. Limiting the collection of information in this way has been regarded as good information management practice for some time and is strongly associated with an organisation's ability to maintain its customer data quality. It is well known that data that is collected but not used, or not connected to a functional area of the organisation and will degrade in quality very quickly (Orr 1996).

However, many organisations have difficulty with identifying precisely what personal information is and is not necessary for their ongoing functions and activities. In addition, there is a prevalent tendency to strive for data quality at the semantic level of completeness by collecting a broad range of personal information about customers (Gibbs et al. 2002). One interviewee stated:

The approach has tended to be 'Let's collect a whole lot, just in case, and then we're pretty well covered if, you know...'

Striving for completeness in customer data can easily and quickly lead to poor data quality across its other dimensions. Remembering that quality is defined as 'fitness for purpose' it is easy to see why this would be the case. Customer data that has no purpose is, by definition, of poor quality (Shanks and Darke 1998). Data collection that is unconnected to a current organisational function tends to be of low quality because there is no motivation to maintain strong quality control at the point of collection. Secondly, data quality problems of these kinds with personal information are typically identified and corrected when that data is actually used such as when a transaction is being completed with the customer.

The challenges faced by organisations with undisciplined collection practices in complying with the provisions of the EUPD and resulting legislation will be to determine the precise purposes that motivate the collection of personal information. Having made this determination, they will need to change their data collection practices and only collect personal information that is required for specific

activities or functions. Thus we can observe that complying with these requirements will have a positive impact on data quality by encouraging good information management practices.

4.3 Controlling Use and Disclosure Across the Organisation

One interviewee, a consultant from a large accountancy firm with a number of years experience in performing privacy audits as well as providing consulting services to public and private sector organizations, made the following observation:

I think some organisations get quite a surprise when they actually look at how they use information and who they disclose it to. Whereas they might think it's all relatively under control, suddenly – you know, you might have a department collecting [personal information] in a structurally diverse organisation with different people sharing across different quarters – it becomes a big issue because, generally, no one knows what everyone else is doing with the information.

The previously stated sections of Articles 6 and 17 relating to use and disclosure require organisations to only use or disclose personal information for the primary or related secondary purposes for which it was collected and/or for which consent has been obtained. In this regard the EUPD is very clear; use of personal information must be restricted to those purposes the individual has been informed about and to which they have consented. However, large organisations will commonly use customer data for a variety of purposes and sharing of this personal information across an organisation's functional units or business lines is often necessary. If different units in an organisation are to use a common customer data set, they must maintain a shared understanding of what is, and is not, an appropriate use or disclosure of the personal information it contains. Data quality at this social level of shared understanding and common interpretation must also be maintained for the total duration that each piece of personal information is stored and used by the organisation; a duration that could easily extend for many years or decades.

Therefore, in order to comply with these articles it will be necessary for organisations to tag personal information held in customer data sets with its approved and allowable uses. This is particularly important when customers are giving consent for certain uses and not for others. It also important when different functional units are using customer data for unrelated purposes. In the Australian context it has been estimated by industry observers that the cost of reprogramming information systems to allow for this tagging costing large retail and financial service firms millions of dollars (Sinclair 2002).

An important further consideration in the EUPD is contained in Article 14 which gives a subject an opportunity to actively prevent disclosure to third parties for purposes such as direct marketing. Maintaining a record of customers that have and have not, opted-out of direct marketing campaigns can pose a number of similar difficulties for large organisations.

Establishing a flag for this purpose in the organisation's customer databases is one viable option provided the technical capacity for an expansion of this nature exists. However, some organisations do not possess this capacity. Also, some large organisations wish to keep this kind of marketing information separate from their customer databases for a variety of reasons. One of our interviewees, the CIO (Chief Information Officer) for a large retail company, reported that his organisation maintained opt-out lists on a PC-based system separate to their mainframe-based customer database and used the opt-out lists to cleanse their customer lists before mounting a direct marketing campaign. Another interviewee who consulted to large companies on privacy issues also suggested that this kind of practice was quite common in industry. These kinds of practices raise a number of issues for the maintenance of customer data quality particularly if this opt-out list is maintained within a single department, such as marketing, and general access is not possible to staff

who have direct customer contact such as front of house and call centre staff. The real risk is that unless this opt-out information can be updated in an easy and timely manner, it will get lost or improperly recorded, with implications for the organisation if individuals who have opted-out of direct marketing campaigns continue to receive this type of material.

4.4 Allowing Anonymity when Interacting with Customers

Where possible, organisations must give individuals the option to remain anonymous during any contact or transaction they have with the organisation. For example, it should be possible for an individual to telephone an insurance company to obtain an estimate for house and contents insurance without having to provide any identifying information such as a name, telephone number or street address although non-identifying information such as postcode and value of goods to be insured might need to be provided.

However, for many organisations, conducting anonymous transactions with members of the public is difficult due to constraints built into their transactional information systems. According to one of our interviewees involved in privacy consultancy work:

...we can't because of our systems constraints. We can't actually [have anonymous transactions], we have to go through certain identification fields before we can provide information.

Many information systems require personal information to be entered into mandatory data fields before a transaction can be processed and information or a similar service can be provided. Thus, in the example above, in order to obtain a quote for insurance, it might be necessary for a call-centre operator to complete a number of data fields with identifying information before the system will provide an insurance estimate. Of course, one common workaround often used to circumvent these kinds of information systems constraints – constraints that have often been built in deliberately in the first place in an attempt to improve data quality by ensuring completeness – is for operators to enter dummy data in order to move through the system to gain the required information. These kinds of practice can seriously degrade customer data quality. While existing customer data quality does not, in itself, directly affect an organisation's ability to comply with Article 6.1.e (anonymity), information systems constraints designed to preserve semantic data quality can impede an organisation's ability and willingness to comply with the anonymity requirements of the EUPD and raises further issues for the way systems should be developed.

5. Implications For Systems Development

A number of important implications for practitioners emerge from this study. First, the EUPD and resulting legislation have been designed to improve the information privacy of individuals by 'giving them some control' over how their personal information is used by private sector organisations. However, in order to 'give control' organisations must have control over this information in the first place. As we have shown in this paper, poor customer data quality severely undermines the ability of organisations to control the personal information they hold about individuals and inhibits their ability to comply with the new legislation.

Second, it would seem that most legislation aimed at protecting information privacy is based on the assumption that organizations have an integrated customer data set and that it is relatively easy to access, collect and collate all the personal information they hold about an individual. The reality is quite different for most organizations. These organisations cannot readily achieve the whole of customer view necessary for strict compliance with the provisions of the EUPD due to problems with their customer data quality.

Third, this problem is particularly pernicious for organisations with multiple points of customer contact. These organisations 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 whole of customer view. This inability to develop an integrated whole of customer view compromises an organisation's ability to effectively manage its customer data and hence compromises its ability to meet its obligations under the EUPD. These organisations will need to exert strong control over the ways in which their function units collect and manage personal information if they are to improve their customer data quality sufficiently to achieve a whole of customer view and comply with the EUPD.

Fourth, the ability to develop a unified and integrated, whole of customer view about an individual enables organisations to unproblematically comply with the provisions of the directive. Establishing and maintaining high levels of customer data quality across all four 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 those kinds of information systems that use good quality, highly integrated databases of personal information that have raised the hackles, suspicions and fears of privacy advocates and political commentators for several decades due to their ability to data mine and match data from multiple sources (See for example Davies 1997, Davies and Hosein 1998). Yet, it is precisely those organisations with highly integrated and carefully managed customer data that are in the best position to comply with the provisions of the new legislation.

6. Conclusion

While many organisations have spent some time and effort on the window dressing of privacy policies and disclaimers, they have done little to change their underlying processes and to redesign their data infrastructure to deal with the required changes to the way in which they handle personal information (Gibbs et al. 2002). Given the issues identified in this study, we believe this signals a significant problem which needs to be dealt with in future information systems development.

Nevertheless, it cannot be denied that the EUPD has advanced information privacy protection in compliant countries by requiring organisations to formulate privacy policies and to take reasonable steps to protect the information privacy of individuals. Privacy issues consistently rank highly in consumers' list of concerns (Clarke 1997, Davies, 1997) and have been attributed as one of the factors limiting the growth of e-commerce (OFPC 2001). Complying with the principles not only prevents possible damage to reputation resulting from non-compliance but can help develop much needed trust between private sector organisations and their customers. In addition, the EUPD has succeeded in raising awareness of information privacy issues amongst IS practitioners as well as in the broader community and has quite visibly started a cultural shift in the private sector towards a culture that respects and values information privacy.

7. References

- Clarke, R. (1997). What Do People Really Think? MasterCard's Survey of the Australian Public's Attitudes to Privacy. *Privacy Law and Policy Report*. 3(8), 141-142.
- Davies, S. (1997) Re-Engineering the Right to Privacy: How Privacy Has Been Transformed from a Right to a Commodity. In *Technology and Privacy: The New Landscape* (Agre, P. and M. Rotenberg Eds), MIT Press, Cambridge Mass.

Davies, S. and I. Hosein (1998). Liberty on the Line. In *Liberating Cyberspace: Civil Liberties, Human Rights, and the Internet* (Cooper, J. Ed.), Pluto Press, London.

EC/95 (1995) Directive 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, 20 Feb.

English, L. (1999). *Improving Data Warehouse and Business Information Quality*. Wiley, New York.

Gavison, R. (1980). Privacy and the Limits of Law, *Yale Law Journal* 89, 421-428.

Gibbs, M., Shanks, G., Lederman, R., de Silva, R. (2002) Privacy and Customer Data Quality: Exploring the Issues. In *Proceedings of the 13th Australasian Conference on Information Systems* (McGrath, M et al. Ed.), Victoria University, Melbourne.

Halford v United Kingdom (1997). (Application No 20605/92), 24 EHRR 523, 25 June.

Kahn, B. et al.. (2002). Information Quality Benchmarks: Product and Service Performance. *Communications of the ACM*, 45(4), 184-192.

Lindland, O. et al. (1994). Understanding Quality in Conceptual Modelling, *IEEE Software*, 11(2), 42-49.

Michael, J (1994) *Privacy and Human Rights: An International and Comparative Study, with Special Reference to Developments in Information Technology*, Dartmouth, Aldershot.

Miles, M.B. and A.M. Huberman (1994). *Qualitative Data Analysis: An Expanded Source Book*. 2nd edition. Sage, Thousand Oaks.

Office of the Federal Privacy Commissioner (OFPC) (2001) Privacy and the Community, AGPS, Canberra. URL <http://www.privacy.gov.au> Accessed 9 March 2002.

Orr, K. (1996). Data Quality and Systems Theory. In *Proceedings of the International Conference on Information Quality* (Wang, R. Ed) MIT, Boston.

Redman, T. (2001). *Data Quality: The Field Guide*. Digital Press, New Jersey.

Shanks, G. and B. Corbitt (1999). Understanding Data Quality: Social and Cultural Aspects. In *Proceedings of the 10th Australasian Conference on Information Systems* (Hope, B and P. Yoong Eds) Victoria University of Wellington, New Zealand.

Shanks, G. and P. Darke (1998) Understanding Data Quality in Data Warehousing: A Semiotic Approach. In *Proceedings of the International Conference on Information Quality* (Chengilar-Smith, I. and L. Pipino Eds), MIT, Boston.

Shanks, G. and E. Tay (2001). The Role of Knowledge Management in Moving to a Customer-focused Organisation. In *Proceedings of the 9th European Conference on Information Systems* (Smithson, S. et al. eds). Moderna, Kranj, Slovenia..

Sinclair, J. (2002). Eyes Wide Open, *The Age*. 9 April, Next Supplement, 7.

Stamper, R. (1992). Signs, Organisations, Norms and Information Systems. In *Proceedings of the 3rd Australian Conference on Information Systems* (McGregor, R. Ed.), University of Wollongong, Wollongong.

Strong, D.M. (1997). Data Quality in Context. *Communications of the ACM*, 40(5) 103-110.

Wand, Y. and R. Wang (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.