

Data Governance

DOI 10.1007/s12599-011-0162-8

The Author

Dr. Boris Otto (✉)
 Institute of Information Management
 University of St. Gallen
 Müller-Friedberg-Strasse 8
 9000 St. Gallen
 Schweiz
boris.otto@unisg.ch

Accepted after two revisions by
 Prof. Dr. Sinz.

Published online: 2011-06-08

This article is also available in German in print and via <http://www.wirtschaftsinformatik.de>: Otto B (2011) Data Governance. WIRTSCHAFTSINFORMATIK. doi: 10.1007/s11576-011-0275-1.

© Gabler Verlag 2011

1 Data Quality as a Success Factor in Enterprises

The quality of data is critical for enterprises in order to be able to meet a variety of business requirements, such as compliance with regulatory and legal provisions, integrated customer management (“360° view on the customer”), effective and efficient reporting (“single point of truth”), or integrated and automated business processes.

Consumer goods manufacturer Nestlé, for example, is confronted with requirements from the French retail industry to provide “carbon footprint” information on the packaging of each product shipped to stores. The carbon footprint is supposed to inform about the carbon dioxide emitted during the production and distribution of the product along the entire supply chain (AFNOR 2009). This information has to be specified as an attribute of the product data class and has to be made available for the production and packaging process correctly, completely, and in a timely manner. Otherwise, the company risks being fined.

Such requirements particularly aggravate the management of corporate data (i.e. data that is used across the entire company) in large enterprises, which are typically characterized by complex and often globally spread organizational structures. Such corporate data are, for example, master data on materials, suppliers, and customers (Loshin 2008, pp. 5 ff.). Data Governance is a possible approach to meet these challenges as it specifies who makes decisions with regard to certain data, and what are the tasks and duties resulting from such decisions. In the case of Nestlé, Data Governance is applied to ensure that the right data source is used for providing correct information on carbon dioxide emissions and to specify time, form, and quality of this data which is supposed to be available for the information being imprinted on the product labels.

2 Definition

A standard definition of the term “Data Governance” can be found neither in the research community nor in the practitioners’ community dealing with information systems. However, proposals defining the term agree that Data Governance refers to the allocation of decision-making rights and related duties in the management of data in enterprises. According to Weber et al. (2009), for example, Data Governance specifies a structural framework for decision-making rights and responsibilities regarding the use of data in an enterprise. Khatri and Brown (2010) see Data Governance as referring to the assignment of decision-making rights with regard to an enterprise’s “data assets”.

Data Governance aims at maximizing the value of data assets in enterprises. Viewing data as an asset goes back to the 1980s, when methods and knowledge regarding the management of physical goods were transferred to the field of managing immaterial goods, like information and data (Horne 1995) for the first time. Today researchers are discussing whether the value of data can and should be determined for financial accounting purposes (Atkinson and McGaughey 2006). Generally, data only has

a value if it is being used. Data’s “fitness for use” is what Wang (1998) considers as data quality. Poor data quality reduces the value of data assets in an enterprise if their utility is low (Even and Shankaranarayanan 2007, p. 80). Thus, enterprises are anxious to maximize data quality.

Maximizing data quality is the aim of data quality management. DAMA International (2009, p. 20) defines data quality management as a function for “measuring, evaluating, improving, and ensuring data’s fitness for use”. Data quality management thereby is a sub-function of data management, which comprises planning, controlling, and provisioning of data assets (DAMA 2009, p. 4).

The relationship between data management and Data Governance is based on a differentiation proposed by the *International Organization for Standardization (ISO)* regarding Governance and Management (ISO/IEC 2008). Following this differentiation, Data Governance represents the leading function of data management as it specifies which decisions need to be made in data management and who makes these decisions. Data management ensures these decisions are made and appropriate action takes place. **Figure 1** summarizes the fundamental concepts related to Data Governance.

3 State of the Art in Research

There is broad consensus among researchers that Data Governance must find answers to three questions (Khatri and Brown 2010; Pierce et al. 2008; Weber et al. 2009):

- What decisions, with regard to corporate data, need to be made on an enterprise wide level?
- Which roles are involved in the decision-making process?
- How are the roles involved in the decision-making process?

Regarding the first question, information systems research has come up with a number of answers. According to Khatri and Brown (2010), Data Governance related decisions refer to some fundamental principles of data management (the use of data standards, for example), data quality requirements

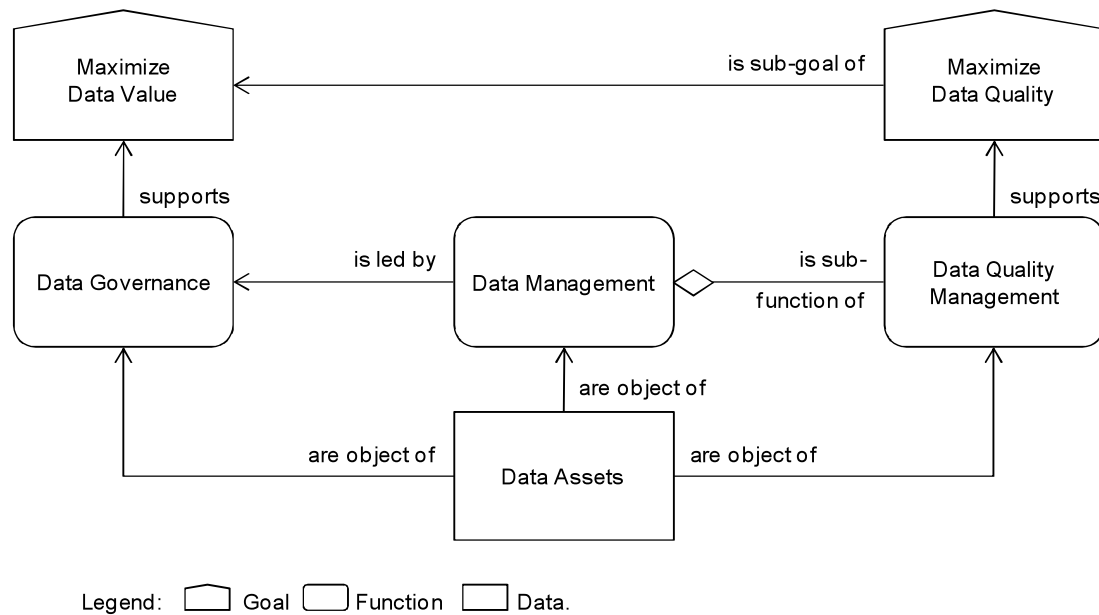


Fig. 1 Fundamental concepts

and data quality measurement, metadata management, data access requirements management, and data lifecycle management. In their differentiated analysis Pierce et al. (2008) identified “company-wide standardization of data definitions”, a “company-wide logical data model”, and “company-wide standardized business rules” to be decision areas of Data Governance.

Similar proposals have come from the practitioner’s community. IBM (2007), for example, considers data quality management, information lifecycle management, and data protection to be core areas of Data Governance, with data architecture management, metadata management, and documentation of review results being supporting tasks.

As for the second question, referring to the roles involved in Data Governance, a number of recommendations can also be found in literature. Roles most frequently mentioned are data stewards, data owners, and data committees. Data stewards support the business departments in the desired use of data (when using standards like eCl@ss for classification of materials, for example). Also, data stewards are responsible for taking up and evaluating business requirements on and problems with data. Data stewards typically deal with data from a certain business department or division (Loshin 2008).

While data stewards represent an enterprise’s data management function, data

owners belong to certain business departments or divisions. They specify the business requirements on the data and the data quality (Khatri and Brown 2010). The term data owner has been criticized by some researchers, as it might suggest that data “is owned” by a certain corporate function or division, which would contradict the approach to view data as company assets (i.e. assets that are owned by the enterprise as a whole). Still, the term data owner has established itself among practitioners. The data owner role regarding supplier master data, for example, is often allocated to the head of central procurement.

A data committee is the central decision-making board in Data Governance. It specifies the principles for using the data throughout the entire enterprise, and it matches the different interests and demands of the functional departments and business divisions (represented by the data owners) on the one side and the data management function (represented by the data stewards) on the other side (Khatri and Brown 2010; Loshin 2008).

Regarding the third question, referring to the linking of roles and decision areas, Data Governance is about assigning authority and – resulting from this – responsibility. For example, the decision-making authority regarding the data architecture could be assigned to the data

committee, whereas the executive power is assigned to the data steward. When assigning roles to decision areas, the basic principle of congruence (in the sense of organizational theory) is to be followed, according to which tasks, responsibility, and authority need to be congruent in order to ensure goal oriented action (Krüger 1994, pp. 47 f.).

To do so, function diagrams are often used for modeling. A function diagram is a technique used in organizational design by which tasks are linked with positions by means of so-called “authority codes” (Schulte-Zurhausen 2005, pp. 515 ff.). In order to design Data Governance for individual enterprises, a number of authors (Loshin 2008, pp. 33 ff.; Weber et al. 2009) propose to use the RACI notation.¹

Besides identifying and describing each of the three elements of Data Governance (decision areas, roles, and authority), researchers are currently investigating the best possible combination of these elements. Khatri and Brown (2010) speak of a “continuum” when assigning decision-making authority to central and/or decentral roles in enterprises. Weber et al. (2009) examined the use of contingency theory for the best possible design of Data Governance under consideration of enterprise specific external and internal parameters.

¹RACI is an acronym made up of the four authority codes *Responsible, Accountable, Consulted* and *Informed*.

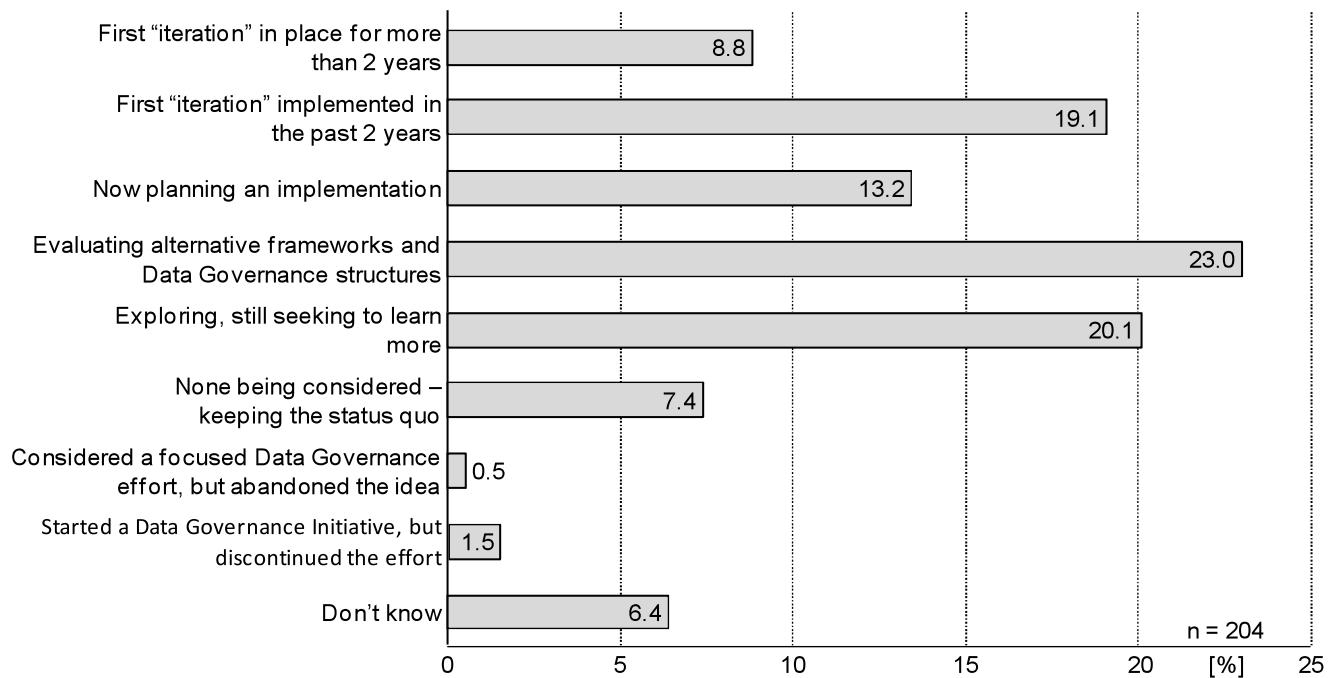


Fig. 2 Data governance state of the art in practice (cf. Pierce et al. 2008, p. 15)

Decision Areas	Roles	BPO	BPE	DST	Data Key User	End User
		global	global	global	local	local
Data Management Organization		A	I	I	I	I
Definition and Standardization of Master Data		A	R	I		
Principles and Guidelines		A	R	I	I	I
Data Maintenance Processes		C	A	R	R	I
Metadata Management and Master Data Documentation			A	R	R	I
Data Quality Measurement			C	A/R	A/R	

Legend: BPO – Business Process Owner; BPE – Business Process Expert; DST – Data Standards Team; A – Accountable; R – Responsible; I – Informed; C – Consulted.

Fig. 3 Data Governance at Ciba (simplified according to Bettschen 2008)

4 Practical Application

Due to the diversity in the manifestation of contingency factors (both enterprise specific like size, degree of diversification, type of decision-making patterns, and industry specific), Data Governance can be found in many variants in practice.

The latest scientific publication on the state of the art of Data Governance in practice comes from Pierce et al. (2008), who conducted a survey among members of the *International Association for Information and Data Quality (IAIDQ)*. The results of this survey reveal a diversified

picture (see Fig. 2). About one fourth of the interviewees (27.9 per cent) said that Data Governance was already in place in their enterprises. The majority of the interviewees (56.3 per cent) said they were currently in the phase of planning Data Governance. 15.8 per cent of the interviewees said they did not intend to establish Data Governance or had dropped plans to establish Data Governance (or they did not provide any information). These findings are confirmed by studies done by software vendors and consulting service providers, like the one by Initiate Systems (Initiate 2010).

An example of an implementation of Data Governance is given by Bettschen (2008). Figure 3 shows the allocation of roles to decision areas by means of authority codes in the case of specialty chemicals company Ciba, which was acquired by BASF in 2009. At Ciba, the data owner role has been assigned to business process owners (BPO) and/or business process experts (BPE). In addition, there are ten data stewards, who are members of the data standards team (DST). The data stewards are assigned to three regional data managers, who report to the

head of the DST. Ciba has not established a data committee.²

As can be seen from the table, Data Governance has to coordinate the interests and demands of different stakeholder groups in an enterprise. Interests and demands need to be matched:

- between business departments and/or divisions (BPO and BPE) and data management (DST);
- between central (global) and distributed (local) organizational units;
- between central standards (DST) and local end users.

5 Future Developments

For further research on Data Governance various topics can be identified. First, a clear distinction between related terms and concepts is required, such as between master data management and data quality management on the one side and “IT Governance” on the other side. Second, evidence is still missing as to whether Data Governance in fact contributes to maximization of the quality of data and, as a consequence, to maximization of the value data has for an enterprise. Third, the question needs to be answered as to whether Data Governance is simply a new

buzzword for concepts already known. For example, in the 1990s Goodhue et al. (1992) had already discussed the possibilities and limits of “Strategic Data Planning”, addressing numerous aspects which are now claimed to be elements of Data Governance (data architecture management, for example). And fourth, answers need to be found as to what effort enterprises should undertake to establish Data Governance without generating too much additional “bureaucracy”, which might counteract the whole endeavor.

References

- AFNOR (2009) General principles for an environmental communication on mass market products. BP X30-323, AFNOR Groupe
- Atkinson K, McGaughey R (2006) Accounting for data: a shortcoming in accounting for intangible assets. *Academy of Accounting and Financial Studies Journal* 10(2):85–95
- Bettschen P (2008) Master data challenges at Ciba. In: 8. CC CDQ workshop, Institute of Information Management, University of St. Gallen
- DAMA (2009) The DAMA guide to the data management body of knowledge. Technics publications, Bradley Beach, NJ
- Even A, Shankaranarayanan G (2007) Utility-driven assessment of data quality. *ACM SIGMIS Database* 38(2):75–93
- Goodhue DL, Kirsch LJ, Quillard JA, Wybo MD (1992) Strategic data planning: lessons from the field. *MIS Quarterly* 16(1):267–274

- Horne NW (1995) Information as an asset – the board agenda. In: *Computer Audit Update* 1995(9):5–11
- IBM (2007) The IBM data governance council maturity model: building a roadmap for effective data governance. IBM Corporation
- Initiate (2010) Data governance survey report <http://www.initiate.com/Documents/Initiate-Data-Governance-Survey-Report.pdf>. Accessed 2010-10-25
- ISO/IEC (2008) ISO/IEC 38500: corporate governance of information technology. ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission)
- Khatri V, Brown CV (2010) Designing data governance. *Communications of the ACM* 53(1):148–152
- Krüger W (1994) *Organisation der Unternehmung*. Kohlhammer, Stuttgart
- Loshin D (2008) *Master data management*. Kaufmann, Burlington
- Pierce E, Dismute WS, Yonke CL (2008) The state of information and data governance – understanding how organizations govern their information and data assets. IAIDQ and UALR-IQ
- Schulte-Zurhausen M (2005) *Organisation*, 4th edn. Vahlen, München
- Wang RY (1998) A product perspective on total data quality management. *Communications of the ACM* 41(2):58–65
- Weber K, Ofner M (2008) Case study Ciba – organizing master data management. BE HSG/CC CDQ/11, Institute of Information Management, University of St. Gallen
- Weber K, Otto B, Österle H (2009) One size does not fit all – a contingency approach to data governance. *ACM Journal of Data and Information Quality* 1:1

²A detailed description on this is given by Weber and Ofner (2008).