



Enterprise Data Management: use and need

Ronald Jonker, Jeroen Tegelaar and Jan Geurtsen

Enterprise Data Management is concerned with the whole spectrum of activities directed toward the organization and proper usage of data. This article covers the background of the relative popularity of data management in many organizations over the past few years. What is the aim of data management? Why is data management so important? And if organizations wish to engage this topic, how to do so? This article strives to answer these questions.



R.A. Jonker
is a partner bij KPMG Management Consulting, IT Advisory.
jonker.ronald@kpmg.nl



J.A.C. Tegelaar
is a senior manager at KPMG Management Consulting, IT Advisory.
tegelaar.jeroen@kpmg.nl



J.M.B. Geurtsen
is a consultant at KPMG Management Consulting, IT Advisory.
geurtsen.jan@kpmg.nl

Introduction

Our digital world is built on information. Data is everywhere and everybody uses it for their everyday business. Data management has recently seen a surge in popularity across companies, organizations, analysts and advisors. What drives this development, considering that this topic is generally not considered to be very sexy?

Data is a representation of facts. By placing data in context, information is created. The absence of good data management often automatically means that management and operational information is not effective. In the long run, this will paralyze an organization to the extent that it can no longer function properly. The reason for this is that important decisions are taken too late, because people can no longer rely on management information. In addition, business process management requires increasing time and effort because tasks and responsibilities of departments are not well aligned to one another. In this situa-

tion, everyone creates and manages only the data that is necessary for the execution of his or her own duties, and accordingly makes use of an own set of data definitions. Because departments accuse one another of inaccuracy, control measures are imposed *en masse* and operations become even more viscous, leading to a downward spiral for the whole venture. Therefore, there is an urgent business reason to assign data management an independent and professional position within organizations.

Moreover, rules and regulations enforce a structured organization of data management. Regulations in the financial sector, such as Basel and Solvency, mandate a governance framework to be implemented for data quality and traceability of information used in organizational reports. In this context, Basel II states: 'The bank must have in place a process for vetting data inputs into a statistical default or loss prediction model which includes an assessment of the accuracy, completeness and appropriateness of the data specific to the assignment of an approved rating.' Qualitative data management is important to be able to meet these criteria, but can also be very time-consuming to implement. In this context, it is difficult to comply with rules and regulations or in place agreements with suppliers and customers.

Enterprise Data Management is comprised of all activities within organizations aimed at the structured identification, classification, registration, modeling, unlocking, securing, archiving and deletion of data. In this frame-

work, the term 'enterprise' represents the organization-wide character of data management.

The fact that data management plays such a crucial role in business operations is underlined by statements from C-level officials. Aloys Kregting, CIO of DSM, chosen as CIO of the Year in 2011, says: 'The CIO should above all be concerned with the value of information. You must know exactly which people need which information when, and facilitate that process as well. This once more underlines the importance of reporting and of master data management.'

As a second example we can point to the CEO of an oil exploration and production company, who realizes that good data management is the next step in his company's progress toward business excellence, and will enable it to stand out from its rivals: 'Continuous improvement efforts will now focus on taking advantage of these changes and uncover the hidden value they offer. This means driving simplified processes and strengthened data management to provide quicker and better-informed decision-making, greater responsiveness to customer needs, and less waste – all resulting in greater competitive performance.'

Data as an asset

As mentioned above, data is a representation of facts. In a business environment this means 'facts concerning business operations'. Without context or structure, this data has no added value to a company. It lacks the content and significance to have any real value. Here, we make a distinction between structured data (stored and arranged in a database) and unstructured data (in the form of documents, files, images, text messages, forms, videos or sound recordings, which cannot be incorporated into rows, columns or records).

Without supplementary information, it is difficult, if not impossible, to classify, register and unlock this data for use. The moment we bring context to this data – that is when it acquires significance. We then add a reference, a date and a time, the significance of the message, a format. With this, the data is structured and becomes information. If we connect all the various sources of information, by establishing relationships and identifying patterns, this information becomes knowledge. This is thus the added value of business intelligence (BI): connecting various information sources in an organization to enhance decision-making by the management of the company. See also Figure 1.

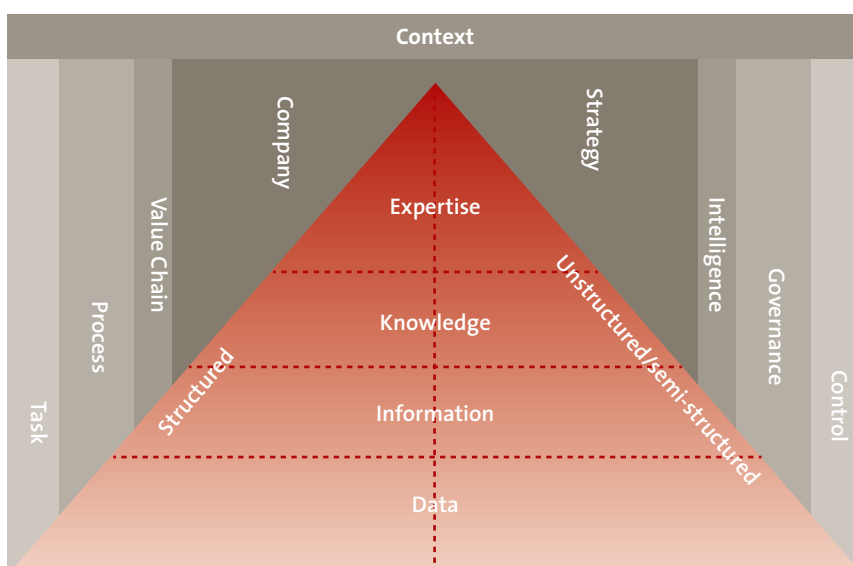


Figure 1. Value of data, placed in context.

Data can be regarded as a company asset

Organizations that are best capable of structuring their data and opening up this information to the knowledge workers within the company will have a competitive advantage. Making use of the inherent commercial power of this data will give companies and organizations a strategic lead over their competitors. Eric Schmidt, erstwhile CEO of Google, stated in 2010: 'I don't believe society understands what happens when everything is available, knowable and recorded by everyone all the time.'¹ And Gartner declares: 'In the private sector we estimate, for example, that a retailer using big data to the full has the potential to increase its operating margin by more than 60%.' ([McK111])

But it's not only about good structuring and unlocking data. For several years, the prevailing idea was that BI would solve the problem of management information. Most global companies and organizations have implemented complex software and executed expensive BI programmes. Nevertheless, management is not satisfied. As BI is primarily oriented toward structured data, insufficient effort is invested in unlocking the value of unstructured data. Moreover, management information cannot easily be modified to accommodate changing company needs. KPMG has stated: 'Huge investments in IT do not necessarily guarantee better information. What is more important is to fundamentally change the way data is gathered, processed and presented.' ([KPMG09])

Information exposed by means of a data warehouse is worthless if the quality of the underlying dataset is poor. Unstructured data (approx. 85% of all company data) cannot be accessed via a data warehouse. The questions are therefore: how can we upgrade this data and what constitutes good data management for unstructured data? In this context, 'good' means in accordance with the quality criteria that the organization has imposed upon the data.² It is evident that 'good data' is not something that simply appears out of thin air. A framework is required. This framework consists of activities that a company must arrange and embed in the organization in a logical and precise manner. This is called data management and covers all organizational activities directed toward business operations, in order to identify, classify, register, model, unlock, secure, archive and delete data in a structured way. For such activities we make use of the term 'Enterprise Data Management' (EDM), because it involves activities that are executed organization-wide.

The awareness that good management of data can add value to company activities and increase profits has brought analysts and advisors to put data on the same level as other company resources such as land, buildings and machinery. In this context, data is defined as a company asset. Assets must be well managed: properly maintained and protected, with assigned ownership and timely disposal or replacement of data if it becomes outdated. Just like other assets, organizational data can also be sold to extract its value. For example, competitors will value customer information because it can be used to improve sales.

Directors of leading companies all over the world have fully recognized this. Data-related programmes are prominent on their action lists. The Hackett Group states: 'What companies are recognizing is that they have thrown lots of money at the applications but, without standardizing and cleansing their data, they are still getting information that does not make sense. They have businesses that are using different definitions, that are calculating metrics differently, that use different hierarchies. This whole concept of master data management is absolutely critical for companies to be able to eventually get to the point where they have predictive analytics.'³ The business case for initiating master data management (MDM) programmes seems evident: 'By 2013, MDM will reduce organizations' data redundancy, which can save 80% of the costs associated with managing redundant data.' ([Gart10])

Models for Enterprise Data Management

The management of data has been the subject of much attention for quite some time, and there is an abundance of models and methods all claiming to provide the best answer to the structure of Enterprise Data Management. The International Organization for Standardization, better known as ISO, has countless standards, each of which covers a sub-aspect of the data spectrum. For example, ISO 27001 deals with information security. ISO 15489 is the norm that is applied to the management of information from an archival perspective. ISO 23081 is the standard for metadata. In addition, one can use ISO 19005 as a guideline for the appearance of data. So we have a glut of standards.

1. See: <http://online.wsj.com/article/SB10001424052748704901104575423294099527212.html>.

2. The article entitled 'Data quality research' by R.A. Jonker in this Compact deals with the notions of quality and quality aspects in more depth.

3. CFO's top priority for 2011? Get a grip on performance management – published on business finance, <http://businessfinancemag.com> – based on a study by the Hackett Group.

Other frameworks such as COSO and frameworks such as Cobit and ISF speak of the importance of data in a wider sense, but only from a risk perspective.

Data Management Body of Knowledge

A more complete model would seem to be that of DAMA-DMBOK. It contains a collection of best practices in the field of data management that have been supplemented by new insights from real-life practice over the years. The DAMA-DMBOK Guide (in full: Data Management Body of Knowledge) is a publication by the Data Management Association, an international organization directed toward data managers and data professionals for the distribution of knowledge about data management.

The DMBOK identifies ten different data functions. These functions are shown in Figure 2. Data governance is the function that links the other domains to one another. In each of the domains, attention should be given to environmental factors, such as current working methods and procedures, techniques used, and the organizational culture.

DAMA does have its weak points. For example, the fact that the functions mentioned only refer to one another in broad terms, meaning that a user does not always recognize or understand the relationship between functions and subsequently the overarching significance of the combination. Moreover, DAMA seems to be oriented toward traditional, structured data, at least at this moment in time. This being the case, little attention is devoted to the

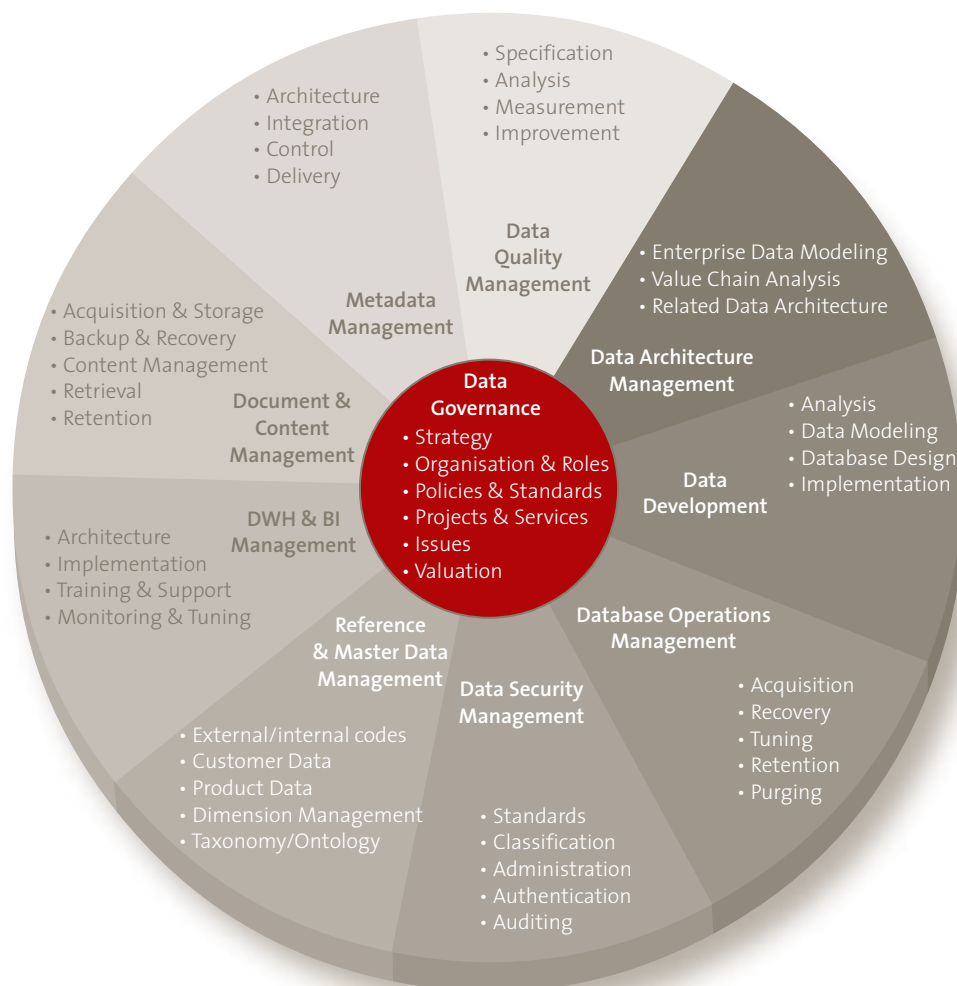


Figure 2. Data domains according to DAMA ([DAMA09]).

importance of content from the social media. Data security within DAMA is primarily aimed at the technological protection of data. Apart from this, the difference in the way generations deal with data has not been explicitly acknowledged as a relevant factor (environmental factor). Finally – and this is perhaps the greatest objection – it is above all a conceptual framework. It lacks practical examples to make concepts and terms sufficiently clear to the reader, entailing a risk of inconsistent interpretation. The way in which the framework ought to be implemented is also rather unclear. This is contrary to the primary goal of a body of knowledge. After all, the application of this body of knowledge should aim at stimulating consistency in the application of data management. It is for these reasons that we use DAMA only for its identification of functions, because those are indeed solid.

KPMG Enterprise Data Management model

The above-mentioned models contain important elements that must be attended to in the realization of a professional data management organization. For the operationalization of data management, however, another set of aspects is also important, aspects that are not covered by these models.

First of all, these involve the fact that data is exchanged between systems both within the organization and between the organization and third parties. Therefore, data management should ensure that good agreements are made about the format in which data is delivered, about validation of the quality of the delivered data, about possible enrichment rounds before the data is further processed, and about any procedures if defects occur in the process. We group these activities under the terms ‘acquisition and authoring’ and ‘distribution’.

In addition, EDM should also ensure that the EDM framework can be maintained as a whole. The organization must have processes at its disposal to record documents and flaws identified during the operational execution of EDM activities. These should be discussed in EDM governance consultation bodies, and should lead to an adjustment of existing procedures and techniques. In this context, one can consider a situation in which a data quality dashboard used within an organization has to be adapted because the organization wishes to monitor a new data object. In such cases, there ought to be a ‘change process’ that sets up the decision-making on this change, and implements the alteration of the dashboard after the decision has been taken.

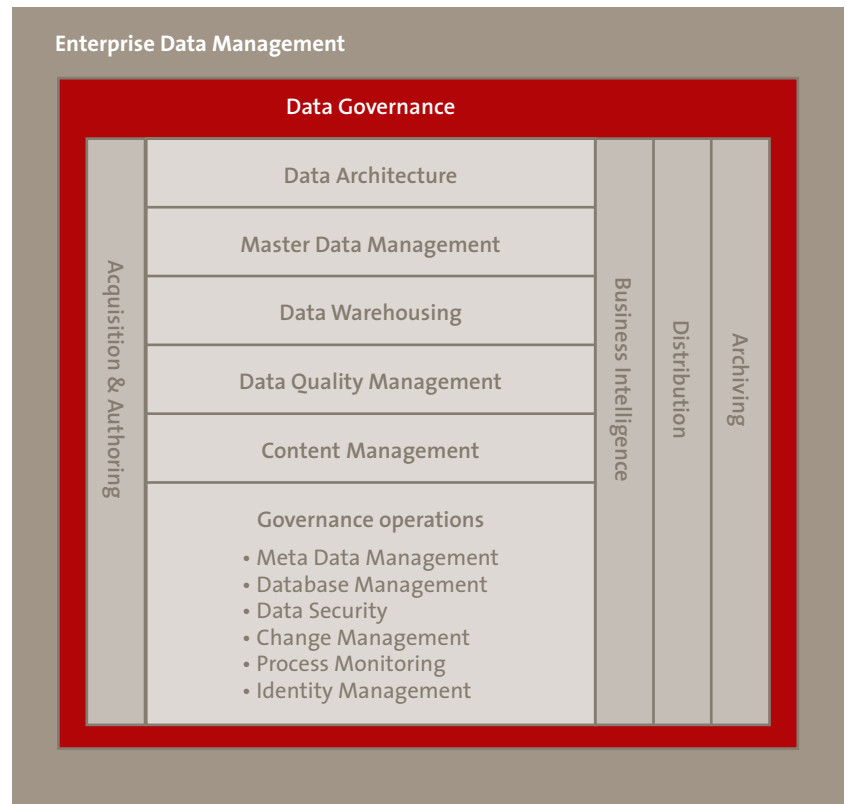


Figure 3. KPMG Enterprise Data Management model.

Finally, all EDM activities performed by an organization should be assessed according to their effectiveness and efficiency. Just as is the case with the primary processes within an organization, there should be a ‘plan, do, check, act’ mechanism for EDM so that one can control whether or not the execution of EDM activities complies with the agreements made on this matter. ‘Process monitoring’ enables this, and allows the EDM organization to independently identify any defects and to take corrective measures.

These steps are depicted in the KPMG EDM model in Figure 3.

A brief description of the most important elements in the model is presented below.

- Data Governance is directed toward the steering of data management activities. Matters such as strategy, policy, roles, tasks and responsibilities come into this category.
- Data Architecture is concerned with the definition and documentation of data objects and data structures in a data model. These form the basis for information analysis and process and system building in an organization.
- Master Data Management concerns the quality of master and reference data. The ultimate goal is to create unique (‘golden’) records.
- Data Warehousing is the activity that ensures the definition of the architecture used to store data in relational databases.

No data management activity whatsoever can be developed and implemented successfully if there is no data governance within the organization

- Business Intelligence involves opening up data that is stored in data warehouses. The data must be provided in such a way that it supplies useful information to management, enabling them to take well-informed decisions.
- Data Quality Management concerns a structural documentation of quality criteria, the analysis of actual data quality, and data quality reporting.
- Content Management is directed toward the classification of data, the structuring of document flows, and access to these.
- Archiving is oriented to the relocation of inactive data to other environments.
- Under Governance Operations, 'meta-data' refers to information on data management elements such as technical and functional descriptions of data objects and data models.
- Database Management is directed toward the operational technical management of databases.
- Data Security is directed toward securing data against unauthorized access and use of this data.
- Identity Management, in conclusion, specifies the access to data.

For a more detailed description of a number of these EDM elements, we refer you to the separate contributions on EDM elements that have been included in this Compact.

EDM from an organizational perspective

Within the scope of this article, it only remains for us to respond to the issue of the best way to implement the EDM model in real-life practice.

If you look closely at the various components of EDM, as shown in Figure 3, you get the impression that there is little logical order in these components. Figure 3 demonstrates that there is no proposed prioritization or phasing of the construction and implementation of the elements. However, it is evident that data governance connects all the other elements. With this, we wish to indicate that there is no ranking between the domains, and that the order of sequence in which the components of EDM are arranged is purely random. Data governance forms an exception here. The link that data governance forms between all the other parts of EDM clearly shows that no data management activity whatsoever can be developed

and implemented successfully if there is no data governance within the organization.

Data governance lays the foundation for all data management activities. Without this foundation, the activities would be merely a pile of loose bricks without structure and cement. This could mean that BI solutions are purchased and implemented while there are insufficient data standards or data definitions. Or it may be the case that the data quality required to generate reliable management information is inadequate. This may lead to the design and purchase of systems that are not compatible with other systems because there is no overarching enterprise data model to serve as the basis for all system developments. It may ultimately result in an organization making active use of the traces that internet users leave behind on websites, without taking privacy rules into consideration, which could lead to image damage and perhaps claims.

Data governance ensures that there is an organization-wide vision and strategy for data management, supported by management. The vision informs us of what we wish to achieve. It indicates the ambition of the organization as it were. All data-related activities should comply to this vision, and strategy ought to bring consistency in these activities. Strategy also dictates the scope of data management within an organization. Ignoring the overarching DAMA model, organizations may prefer to omit certain aspects from consideration because they are probably already being filled in somewhere else, in a decentral unit. A consistently recurring phenomenon, for example, is the fact that HR creates its own data management organization and makes only limited use of the guidelines and standards that the central data management organization has developed.

Data governance also ensures that attention is devoted to the formulation of policy rules. In this context, we are referring to information security policy, policy rules concerning data architecture, archiving and data quality. In addition, data governance ensures the organizational embedding of data management. It is necessary to determine: who is ultimately responsible, where and how are

decisions made on strategy, policy, standards, roles, ownership? For example, how and when are reports on data management activities within the organization formulated? In which way do we organize the execution of master data maintenance activities?

This overview will have made it clear that data governance is the basis of good data management. Regardless of the stage of maturity in which an organization may find itself, it is always beneficial to seriously examine the quality of data governance and to check whether or not its reach is adequate.

Imagine that an organization has its data governance completely in order. Are there then footholds available or best practices that can clarify which of the other data management components are directly eligible for optimization, in terms of prioritization? Unfortunately, this is not the case. In other words, experience has taught us that this depends on the priorities that issue from the agenda of the organization itself.

Imagine that an organization decides to replace a legacy information system with a new ERP system. One might then wonder about the impact this could have on data management. What should have the highest priority? This may lead to 'Data Quality Management' being assigned highest priority as a consequence of the necessary migration. Polluted data is cleansed, meta-documentation is tackled, and the master data management is improved. The implementation of a data integration application may lead, for example, to the data architecture model being updated and a data quality application being selected and

implemented in order to cleanse and enrich data before it is shared with other platforms.

Concluding, we believe that, on a basis of data governance and depending on the business agenda of the organization, those data management activities that bring the most added value in the realization of the agenda at a particular moment should be pursued. The details are shown in Figure 4. Centered on vision and strategy, the business model needed to realize the objectives declared in the vision and strategy is constructed. This business model makes demands on the primary and supporting processes. Resources are needed to enable these processes to function, and can later be subdivided into manpower, data and IT resources. Exactly what and how much is needed on the data side in a specific case is determined by the business agenda. EDM offers a foothold for the way in which this should be organized. This comprises a tailor-made approach and cannot be encapsulated in a fixed pattern of data management activities.

Conclusion

In this contribution we have given an introduction to EDM as an approach to the management of all the data an organization generates or acquires. A proper implementation of this approach ensures that this data complies with the organization's data quality requirements, and that the data needed to execute processes and to enable management to take well-founded decisions is correct, complete, and available timely. When this is the case, data is an asset that must be managed just like all other company assets.

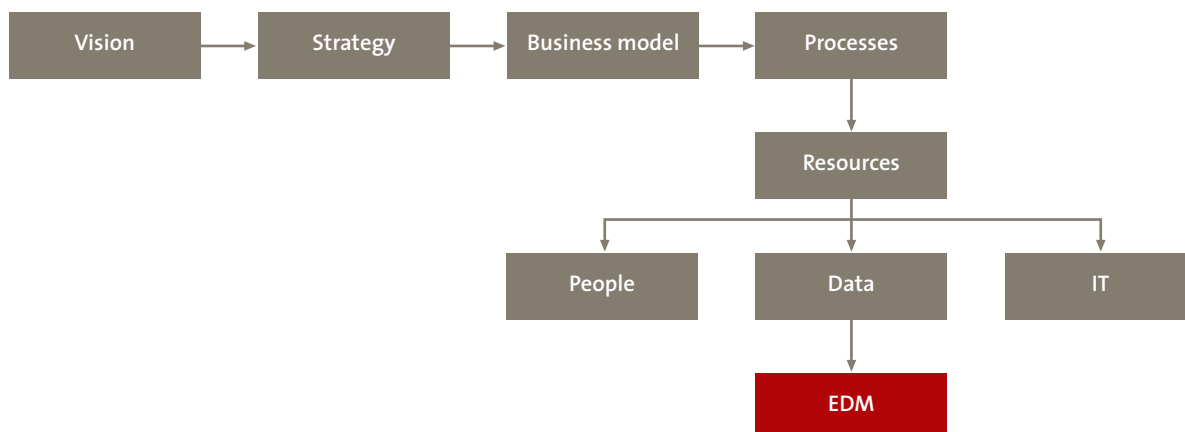


Figure 4. Relationship between business model and EDM.

In the operationalization, it is the company strategy and prioritization that determine which of the components of EDM are selected

Subsequently, we have further defined the constituent parts of EDM. Thus, a framework of management activities has arisen that form the basis for data quality. Finally, we have argued that the implementation of the constituent parts cannot take place according to a fixed pattern. In the operationalization, it is the company strategy and prioritization that determine which of the components of EDM are selected and optimized. A crucial role is allocated to data governance, which ensures the organization-wide and management-sponsored vision and strategy.

References

- [DAMA1] *The DAMA Guide to The Data management Body of Knowledge* (DAMA-DMBOK Guide), p. 7. First edition, 2009. Via <http://franklybi.blogspot.com/>.
- [Gart10] Gartner, *Hype Cycle for Master Data Management*, 2010.
- [KPMG09] KPMG International, *Does your Business Intelligence Tell You the Whole Story?*, 2009.
- [McKin11] McKinsey Global Institute, *Big Data: the Next Frontier for Innovation, Competition and Productivity*, McKinsey & Company, 2011.

About the authors

- R.A. Jonker** is a partner at KPMG Management Consulting, IT Advisory and is a Service Leader for Enterprise Data Management. He advises companies and organizations on the formulation and implementation of organization-wide data management applications.
- J.A.C. Tegelaar** is a senior manager at KPMG Management Consulting, IT Advisory. Within the Service Enterprise Data Management he is the point of address for questions in the domain of data governance and content management. His further specialisms are issues in the field of the management and accessibility of unstructured data, and issues in the field of archiving.
- J.M.B. Geurtsen** is a consultant at KPMG Management Consulting, IT Advisory. In the past few years he has been engaged in processes of information protection and data archiving, particularly in the financial sector.