

CEPIS UPGRADE is the European Journal for the Informatics Professional, published bi-monthly at <<http://cepis.org/upgrade>>

Publisher

CEPIS UPGRADE is published by CEPIS (Council of European Professional Informatics Societies, <<http://www.cepis.org/>>), in cooperation with the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*, <<http://www.ati.es/>>) and its journal *Novática*

CEPIS UPGRADE monographs are published jointly with *Novática*, that publishes them in Spanish (full version printed; summary, abstracts and some articles online)

CEPIS UPGRADE was created in October 2000 by CEPIS and was first published by *Novática* and **INFORMATIK/INFORMATIQUE**, bimonthly journal of **SVI/FSI** (Swiss Federation of Professional Informatics Societies)

CEPIS UPGRADE is the anchor point for UPENET (UPGRADE European NETWORK), the network of CEPIS member societies' publications, that currently includes the following ones:

- **inforeview**, magazine from the Serbian CEPIS society JISA
- **Informatica**, journal from the Slovenian CEPIS society SDI
- **Informatik-Spektrum**, journal published by Springer Verlag on behalf of the CEPIS societies GI, Germany, and SI, Switzerland
- **ITNOW**, magazine published by Oxford University Press on behalf of the British CEPIS society BCS
- **Mondo Digitale**, digital journal from the Italian CEPIS society AICA
- **Novática**, journal from the Spanish CEPIS society ATI
- **OCG Journal**, journal from the Austrian CEPIS society OCG
- **Pliroforiki**, journal from the Cyprus CEPIS society CCS
- **Tölvumál**, journal from the Icelandic CEPIS society ISIP

Editorial Team

Chief Editor: Llorenç Pagés-Casas

Deputy Chief Editor: Rafael Fernández Calvo

Associate Editor: Fiona Fanning

Editorial Board

Prof. Vasile Baltac, CEPIS President

Prof. Wolfried Stucky, CEPIS Former President

Prof. Nello Scarabottolo, CEPIS President Elect

Luis Fernández-Sanz, ATI (Spain)

Llorenç Pagés-Casas, ATI (Spain)

François Louis Nicolet, SI (Switzerland)

Roberto Carniel, ALSI – Tecnoteca (Italy)

UPENET Advisory Board

Dubravka Dukic (inforeview, Serbia)

Matjaz Gams (Informatica, Slovenia)

Hermann Engesser (Informatik-Spektrum, Germany and Switzerland)

Brian Runciman (ITNOW, United Kingdom)

Franco Filippazzi (Mondo Digitale, Italy)

Llorenç Pagés-Casas (Novática, Spain)

Veith Risak (OCG Journal, Austria)

Panicos Masouras (Pliroforiki, Cyprus)

Thorvardur Kári Ólafsson (Tölvumál, Iceland)

Rafael Fernández Calvo (Coordination)

English Language Editors: Mike Andersson, David Cash, Arthur Cook, Tracey Darch, Laura Davies, Nick Dunn, Rodney Fennemore, Hilary Green, Roger Harris, Jim Holder, Pat Moody.

Cover page designed by Concha Arias-Pérez

"Upcoming Resolution" / © ATI 2011

Layout Design: François Louis Nicolet

Composition: Jorge Lácer-Gil de Rames

Editorial correspondence: Llorenç Pagés-Casas <pages@ati.es>

Advertising correspondence: <info@cepis.org>

Subscriptions

If you wish to subscribe to CEPIS UPGRADE please send an email to info@cepis.org with 'Subscribe to UPGRADE' as the subject of the email or follow the link 'Subscribe to UPGRADE' at <<http://www.cepis.org/upgrade>>

Copyright

© Novática 2011 (for the monograph)

© CEPIS 2011 (for the sections Editorial, UPENET and CEPIS News)

All rights reserved under otherwise stated. Abstracting is permitted with credit to the source. For copying, reprint, or republication permission, contact the Editorial Team

The opinions expressed by the authors are their exclusive responsibility

ISSN 1684-5285

Monograph of next issue (October 2011)

"Green ICT"

(The full schedule of CEPIS UPGRADE is available at our website)



The European Journal for the Informatics Professional

<http://cepis.org/upgrade>

Vol. XII, issue No. 3, July 2011

Monograph

Business Intelligence

(published jointly with *Novática**)

Guest Editors: *Jorge Fernández-González and Mouhib Alnoukari*

- 2 Presentation. Business Intelligence: Improving Decision-Making in Organizations — *Jorge Fernández-González and Mouhib Alnoukari*
- 4 Business Information Visualization — *Josep-Lluís Cano-Giner*
- 14 BI Usability: Evolution and Tendencies — *R. Dario Bernabeu and Mariano A. García-Mattío*
- 20 Towards Business Intelligence Maturity — *Paul Hawking*
- 29 Business Intelligence Solutions: Choosing the Best solution for your Organization — *Mahmoud Alnahlawi*
- 38 Strategic Business Intelligence for NGOs — *Diego Arenas-Contreras*
- 43 Data Governance, what? how? why? — *Óscar Alonso-Llombart*
- 49 Designing Data Integration: The ETL Pattern Approach — *Veit Köppen, Björn Brüggemann, and Bettina Berendt*
- 56 Business Intelligence and Agile Methodologies for Knowledge-Based Organizations: Cross-Disciplinary Applications — *Mouhib Alnoukari*
- 60 Social Networks for Business Intelligence — *Marie-Aude Aufaure and Etienne Cuvelier*

UPENET (UPGRADE European NETWORK)

- 67 From **Novática** (ATI, Spain)
Free Software
AVBOT: Detecting and fixing Vandalism in Wikipedia — *Emilio-José Rodríguez-Posada* — Winner of the 5th Edition of the *Novática* Award
- 71 From **Pliroforiki** (CCS, Cyprus)
Enterprise Information Systems
Critical Success Factors for the Implementation of an Enterprise Resource Planning System — *Kyriaki Georgiou and Kyriakos E. Georgiou*

CEPIS NEWS

- 77 Selected CEPIS News — *Fiona Fanning*

* This monograph will be also published in Spanish (full version printed; summary, abstracts, and some articles online) by *Novática*, journal of the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*) at <<http://www.ati.es/novatica/>>.

Data Governance, what? how? why?

Óscar Alonso-Llombart

Decision-making is based on the information we obtain from business data. All decision-making involves accepting a certain degree of risk, but the truth is that it is not always possible to have complete and hard data available ... In this situation, how can we achieve real value from the data we do have and provide a consistent view of the business performance?, how can we properly analyse the available data taking into account the constant changes that take place in every organisation?

Keywords: Data Governance, Data Management, Data Ownership, Data Quality, Data Stewardship.

1 Introduction

Being able to obtain the real value of data is not an easy task. Data can be collected from multiple channels and then stored in different information systems and databases hosted on heterogeneous technology environments and in differing formats. Even when we have direct access to data, it is difficult to make use of them where, when and how we need to. Also data are often "dirty", full of errors, omissions and/or inconsistencies.

This issue is important enough to make any ICT strategic business project, initiative or even an entire company fail miserably. The data layer of an organization is a critical component, with which overly optimistic assumptions are often made and the real quality of the data is misunderstood or even ignored.

Data is used only in a technological environment is usually restricted to a process or an application with limited impact. There is also some data which is critical because it defines the most important identities (customers, products, employees, suppliers ...) and this has to be shared by multiple processes, departments and business lines. This data (called "master data") should be treated as a strategic asset.

Ensuring quality, integrity and accuracy of data is one of our greatest challenges. Ensuring a clear and consistent view of data across departments, lines of business or other groupings in a modern company, can be critical to the achievement of business objectives.

Achieving quality data is a philosophy that aligns strategy, business culture, and technology to manage data for the overall benefit of the company. In short, this is a competitive strategy that every company can use to differentiate themselves from their competitors through their data quality and their use of data.

“Being able to obtain the real value of data is not an easy task ... data are often "dirty", full of errors, omissions and/or inconsistencies”

Author

Óscar Alonso-Llombart is Engineer in Management Software by the *Universitat Autònoma de Barcelona*, Spain, has a Master's degree in Software Engineering by the *Universitat Politècnica de Catalunya*, Spain, and is Graduate in Data Mining by the *Universitat Oberta de Catalunya*. He works as Analysis Manager at the Spanish company Penteo. He has over 15 years of experience in technology consulting in areas such as Business Intelligence, Datawarehousing, Corporate Performance Management, custom development, implementation of development methodologies, etc. He is author of numerous articles and studies on the application of information systems to the business strategies. Twitter: <@oalonsollombart>; LinkedIn: <<http://www.linkedin.com/in/oscaralonsollombart>>. <o.alonso@penteo.com>

But to what extent does poor data quality affect today's business? Due to the dynamic nature of data, which is typically generated by numerous business processes and combined information sources, stored and used in various systems, it is a major challenge to establish methods to evaluate the impact of poor data quality.

Despite this challenge, it is clear that poor data quality has a real economic cost, primarily in the efficiency of processes.

From research conducted by Penteo it is clear that there is still a considerable *gap* between today's insight and true business intelligence for almost all companies. Although there are many companies that have implemented business intelligence systems, a significant percentage have only done so in isolated projects, responding to very specific needs. In the vast majority of the companies the challenge is to find and properly exploit the data and information on status and progress of the business itself (see Figure 1). This situation invariably impacts the business in terms of economics, confidence about the data, regulatory compliance, satisfaction and productivity.

2 Data Management

Business processes rely heavily on information systems, systems that interact with each other, sharing information and being able to communicate in order to provide adequate and efficient service to the organization. It is possible to make strategic decisions based on information extracted from the systems, and we must have reliable information

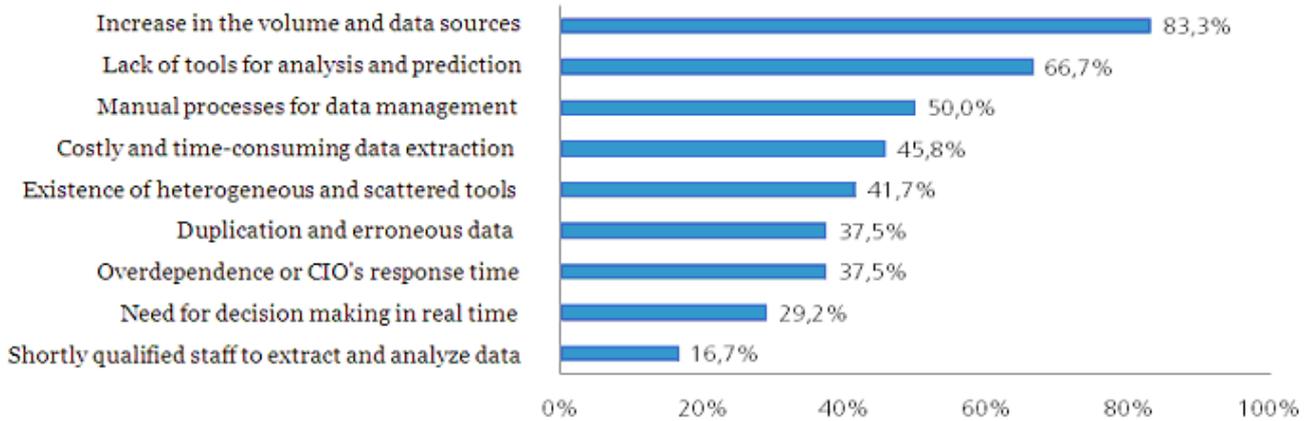


Figure 1: What are the Main Problems in making Decisions? [Source: Penteo.]

for a good corporate management.

In this situation we must realize that we are dependent on the quality of the data we have in our organization. Data as an entity in itself does not add value to business and business intelligence solutions are nothing if we do not have reliable data. Those companies that manage data quality effectively tend to avoid the problems arising from decisions based on poor or misleading information.

Data Management is the first piece on which to base an appropriate use of information (see Figure 2), after considering the data and information inferred from them as valuable business assets. The data and information must be carefully managed, like any other asset, ensuring quality, safety, integrity, availability and effective use.

The objectives of Data Management are:

- To understand the information needs of the organization.

- To capture, store, protect and ensure the integrity of data assets.

- To continuously improving the quality of data and information including accuracy, integrity, integration, relevance and usefulness of data.

- To ensure privacy and confidentiality, and prevent unauthorized and inappropriate use of data and information.

- To maximize effective use and value of data assets and information.

- To be aware of and control the cost of Data Management.

- To promote usage of and deeper and broader knowledge of the value of data assets.

- To manage information in a consistent manner throughout the organization.

- Align Data Management and the technology needed to business needs.

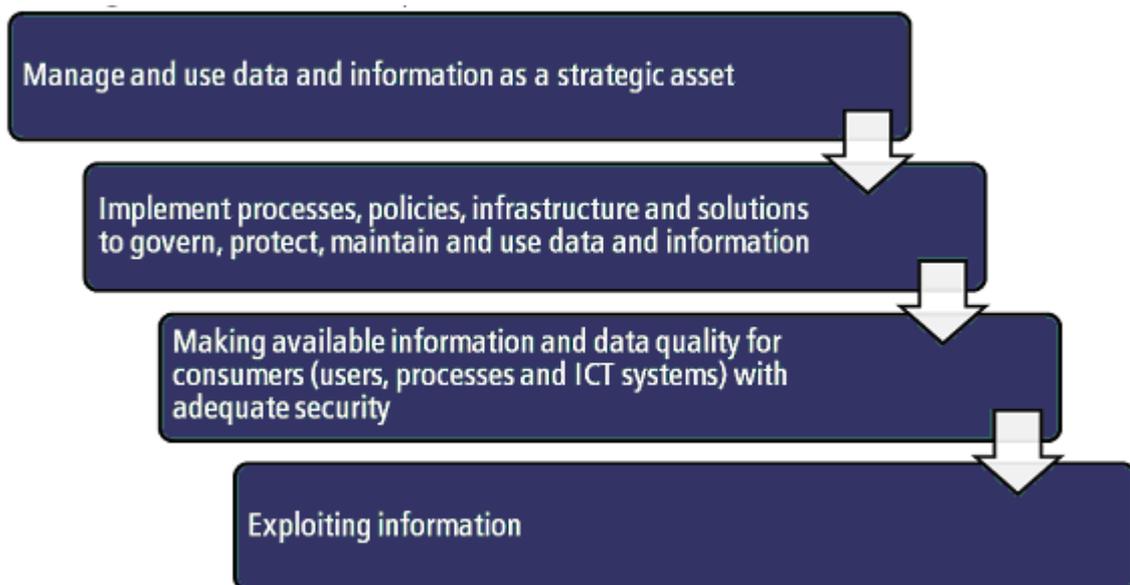


Figure 2: Data Management and Information. [Source: Penteo.]

“ Ensuring quality, integrity and accuracy of data is one of our greatest challenges ”

Data Management has to be seen as a business function. Real competitive advantage is obtained through the appropriate use of the information.

3 Data Governance (Technology cannot solve the Problem by itself)

Data Governance... what is it and why is it important? What is the relationship between governance and ownership of data? Is the concept of Data Management included in data governance? Do we know the costs to the organization of having duplicate data or having no standard definitions of common data? If we are unable to answer these questions, perhaps we should consider a strategy for addressing the need to understand and use data more effectively and efficiently.

To achieve this goal companies must implement Data Governance projects, a set of policies and procedures that, combined, establish the processes that will allow you to transform data into a strategic asset and take the company to a higher level of "maturity" in the use of information, improve data quality and resolve any inconsistencies, managing change in relation to the use of data, and meet regulations and internal and external standards.

Data Governance is the cornerstone on which all practices related to Data Management underpin, that interact and influence each and every one of these, such as data quality, data integration or *warehousing* projects. Data governance is the exercise of authority and control (planning, monitoring and enforcement) on the management of data assets; data do not rule directly but governs how users access data through technology.

A Data Governance program provides guidance on how the other functions of Data Management should operate, appointing data owners on both executive and operational levels. It also has to properly balance objectives with compliance, which limit access to data, and integration of the business processes that increase access to them. The tasks that a Data Governance Program must carry out are:

- Guide information managers in making decisions.
- Ensure that information is consistently defined and understood by all stakeholders.
- Increase the use and reliability of the data as a valuable asset.
- Improve the consistency of projects across the organization.
- Ensure compliance with internal and external regulations.
- Eliminate possible risks associated with use of the data.

Data Governance implementation projects programmes are as unique as the companies that implement them. However, the structural frameworks that are used are actually quite similar to each other. There are common foundational components on which to build the initiative:

- *Organization*: structure responsible for deploying capacity of resources and administration of activities.
- *Policies*: principles and standards, guidelines for information management, and principles to ensure data standards and procedures of government.
- *Processes and practices*: establishing the principles that guide how the policies and processes are created, modified and implemented.
- *Metrics*: a measure to monitor the performance of the government initiative and the actions to significantly improve continues the quality of the data.
- *Data architecture*: including corporate standards data, metadata dictionary, and also security and privacy measures.
- *Tools and technology*: the tasks should be automated using software whenever possible, using data quality tools, data profiling, metaData Management tools, dashboards, etc.

4 Organizing a Team of Data Governance

This is an initiative that should not be considered as an ICT project, but as a continuous process of change in corporate culture. The business must lead the initiative, the implementation of Data Governance is an important change in mindset that must permeate all areas of the company.

Shared responsibility is the hallmark of Data Governance. It requires working across organizational boundaries and systems. Some decisions are primarily a business with input and guidance for ICT, while others are technical decisions and guidelines with input from users at different levels.

The different business units are represented in the "owners" of the data, while Department ICT provides the structure and processes. These data owners, experts in certain subject areas, are put forward as representatives of business interests with the data and take responsibility for the quality and use of these.

If, prior to the implementation of the data governance initiative, there have been Business Intelligence projects it is very possible that there is some sort of Data Governance team. This, although in an informal basis, should help mitigate the costs and organizational changes often required by this type of initiative, and will facilitate us having people who can occupy the profiles that are needed.

“ Poor data quality has a real economic cost, primarily in the efficiency of processes ”



Figure 3: Organizational Chart of the Data Governance Team. [Source: Penteo.]

Staff forming part of the Data Governance team must know how to use and analyse information to facilitate decision-making, and require a mix of technical, analytical and business skills:

- Know the business, its processes, the analytical capabilities of the systems and the company's strategy to establish a master plan for data governance.
- Understand the organization and culture and channel access to information.
- Keep abreast of new capabilities that the technology can bring to the organization.

One of the historical problems in the implementation projects of Data Governance is the lack of adequate monitoring. While some organizations have successfully defined policies and government processes, often they have not put in place the necessary organisational structures to make it work properly.

The organisational framework government data programme must support the needs of all participants throughout the company. With the proper executive support, the Data Governance programme will benefit from the company's participation in the various functions required. This includes both strategic, such as data owners, and tactical, such as coordinators of data teams.

The specific roles include (see Figure 3):

- Director of Data Governance, responsible for managing the initiative and ensuring maximum adoption in the organization. This profile supports the executive sponsors

and provides periodic reports project performance, as well as negotiating with external suppliers of data the associated service level agreements.

- Data Governance Committee, typically multifunctional strategic committee composed of the executive sponsor, the director of the Office of Data Governance, and the CIO of the company. Ideally, executive sponsorship should come from the business area rather than the ICT department. This committee reviews and approves the policies, processes and procedures, managing priorities and evaluates their proper discharge.

- Data coordination team, tactical team that ensures data quality meets the expectations of clients and manages the initiative among the various business units. It is the responsibility of this team to detect and communicate opportunities to the Committee on Data Governance.

- The owners of the data, which manage the lifecycle of data and provide support to the user community. These owners define the criteria for data quality to meet the expectations of the business units, and report the activities and problems with coordination team data.

5 The Need to establish Data Ownership

One of the key factors of successful implementations of Data Governance initiatives is the role of data stewardship or "Owner of the data." The ownership of data is the formalisation of responsibilities to ensure control and effective use of data assets.

“ There is still a considerable *gap* between today's insight and true business intelligence for almost all companies ”

“ Data Management is the first piece on which to base an appropriate use of information ”

The Data owners are business users, experts in specific subject areas designated as responsible for managing data on behalf of other users. They represent the interests of all stakeholders, including, but not limited to, the interests of their own functional areas and departments, protecting, managing and reusing data resources.

These profiles should be a business perspective to ensure quality and effective use of organisational data. The governance process will involve data owners as participants, but they will still be directly responsible for the successful management of data in their domain.

In practice there is no "silver bullet" model that fits all organizations. Basically there are five models of data ownership that organizations can apply, each of these models is unique, with its own pros and cons:

- Model 1: Property subject areas. In this model each data owner runs a particular subject area, as well as the responsibility of the customer data is different from those responsible for product data, etc. In large or complex environments, there may be more than one owner for each subject area. This model works well for companies with multiple departments to share the same data.

- Model 2: property business functions. In this case the owner of the data focuses on data that a department or line of business uses, such as data related to marketing, finance, sales, etc. Depending on the size of the organisation and management complexity, it may be that there are other owners of data by subject area, resulting in a hybrid model with the previous model.

- Model 3: property for business processes. Each business process is assigned a data controller, in this case the data owners are responsible across multiple domains of data or applications involved in a particular business process. This is a very effective model for companies with a clear orientation and a very clear definition of business processes. In organisations where there is no culture of immature processes then this approach is not the best choice.

- Model 4: Property for ICT systems. Those responsible for the data are assigned applications that generate the data they use. This model is a way to evangelize the concept of ownership of the data from the ICT department to the various business units. The data owners can report the progress of the initiative and show how the data will not only improve over time, but also will affect business results.

- Model 5: property projects. Associating the concept of data ownership with projects is a quick and practical way of introducing the culture of Data Management into the or-

ganisation. Unlike the models discussed above, this is a temporary measure which is often used as a starting point for the formal establishment of another long-term model.

To decide the ownership model of ideal data for the organisation is not a trivial task and is one in which we must consider a number of factors such as:

- Profiles and skills available in the organization for Data Management.

- The culture of the company.

- The reputation of the quality of the data.

- The current situation regarding ownership of data.

- Current use of metrics associated with data quality.

- The needs for data reuse.

6 How to tackle the Project of Data Governance?

A proper implementation of Data Governance can have a very positive direct impact on business performance. However, it is a challenge to achieve the right mix of people, processes and technologies to design a successful initiative.

To meet this challenge we must build a data governance strategy effectively, led by business objectives, providing stakeholders with improved capabilities for decision making and helping the company achieve its desired objectives. An effective strategy must ensure that company objectives, business strategy, investment and Data Governance systems are aligned.

A Data Governance initiative is nothing if not driven by the objectives of the company. Business requirements and business objectives should drive the iterations of the project. We need to establish a strategy before introducing the technology into the process.

Before beginning any work with data governance strategy, it is essential to understand and document the overall objectives to help formulate the vision and mission of government data for business growth. After documenting the initial list of objectives the major stakeholders must work to confirm the validity of the list of goals and proper prioritisation. This will ensure that we begin to build our strategy of Data Governance with a suitable base aligned with the business and users.

From Penteo's market analysis and best practices we can draw the following:

- *Engaging business to lead the initiative.* Data Governance is not just a technology but also an important change in mentality that must transcend all areas of the company. Effective change management and communications from the start of the project are essential to ensure success. The

“ Data Governance is the cornerstone on which all practices related to Data Management underpin ”

“ Shared responsibility is the hallmark of Data Governance. It requires working across organizational boundaries and systems ”

project must be addressed from the component organisation and processes, but closely monitored by the ICT department. The historical existence of the role of organisation is emerging as a clear enabler of the adoption of the initiative.

■ *Selling the process internally.* Deployments of Data Governance pose a significant impact on the organisation in many ways, so company CIOs should only start their data governance projects when they have reached a consensus on the decision with other officers involved in the process and when they have managed to successfully sell the project internally. Thus, the involvement in the project from different business areas is fully secured in advance and therefore the risk to address the process is much more controlled.

■ *Adopting Data Governance must not be approached as a finite project.* The change of mentality, culture and the reorientation of the company together with the quality of information are indicators that identify the success of an initiative, hence it cannot be treated as a typical ICT project.

■ *Manage a portfolio of strategic suppliers.* The current market situation forces us to evaluate, monitor and manage the ecosystem of our applications and road map a portfolio of solution providers to standardise and reduce risk, redundancy and cost. The selection of tools has less to do with the features and more with their ability to meet specific business requirements.

■ *Planning and design prior to implementation.* This is a major initiative of high complexity so time must be taken to define exactly the foundation of the future service-oriented system.

Finally, it is important that a Data Governance strategy should be designed to be agile and adaptive. It must be treated as a living entity that is constantly evolving to meet business objectives. The strategy should focus on communicating what is being planned to implement, how it will be implemented and when users will see their needs reflected in the system. Begin with general policies and guidelines and high-level diagrams as the ecosystem will mature in parallel with formal documentation and the level of detail identified in the strategy. It must be ensured that the data governance strategy evolves as part of the vision of the company as the iterative process produces more and more detail. Continuous evaluation and reinvention must be undertaken as business needs change, taking into account the current and future technology trends to support in building and delivering successful data governance strategy.

7 Conclusions

The tangible assets of organisations have a clear value and are managed through information systems and business proc-

esses. The associated data, precisely because of its intangible nature, is not collected on many occasions as strategic assets. However, accurate and available data is a pre-requisite for operations of any organisation to be effective.

Companies that are able to recognise the real value of the data, i.e.: that have established processes, policies and procedures for data quality, are aware of what data is really important or relevant to their business and ultimately rely on the quality of their data, they have become "data-driven organizations." These organizations have an obvious advantage over their competitors by managing the data as a more strategic asset, but to achieve this goal there must be an appropriate strategic vision to improve the quality of information.

The implementation of a Data Governance project requires the support of all business areas involved. Taking control of the data leads to better customer retention, increasing the success of marketing strategies, better control risks and, ultimately, allowing the company to be managed more effectively and efficiently.

Proper implementation of Data Governance eliminates discrepancies between data silos. However, those companies that have implemented these projects have realised at once that the timing of implementation varies greatly depending on the scope and simple exercises that are not technological in nature.

When taken correctly, Data Governance is a discipline helping to achieve the true value of analytic applications and should become the foundation for all initiatives in information management. To achieve proper management of these entities there must exist an appropriate strategic vision to improve the quality of information.

Are those projects that focus iteratively, starting with the set of needs and data that provide the greatest value to the business in the shortest possible time the most successful? Are you looking for a better decision making through Business Systems Intelligence? If the answer is yes to these questions then our starting point must be the analytical data. If we instead seeking to achieve greater operational efficiency or to gain consistency in processes across different transactional systems then we should start with the operational data.

Bibliography

- J. Dyche. Five Models for Data Stewardship. Baseline CONSULTING, 2009.
- D. Loshin. Data Governance for Master Data Management and Beyond. DataFlux, 2008.
- O. Alonso. Trends in the use of BI in Spain 2009. Penteo, 2009.
- O. Alonso. The problem of Data Management. Penteo, 2010.