O UNIVERSITY OF OREGON
APPLIED INFORMATION MANAGEMENT

Applying Aspects of Data Governance from the Private Sector to Public Higher Education

CAPSTONE REPORT

Marsha Poor
HRIS Specialist
Mount Wachusett Community College

University of Oregon
Applied Information
Management
Program

July 2011

Approved by

_____
Dr. Linda F. Ettinger
Senior Academic Director, AIM Program

Applying Aspects of Data Governance from the Private Sector to Public Higher Education

Marsha Poor

Mount Wachusett Community College

Abstract

The purpose of this annotated bibliography is to identify ways to better manage enterprise-wide data assets within institutions of higher education through data quality actions and data governance options. The goal is to present selected data governance practices within the private business sector for consideration by individuals in public higher education who promote and support data quality initiatives. Topics include data quality barriers, data quality models, data quality management practices, and data quality drivers.

Keywords: Data governance; data quality; data steward; high quality data; poor data quality; data asset; higher education

Table of Contents

List of Tables and Figures

Tables

Figures

Introduction to the Annotated Bibliography

Problem

As a Human Resource Information System (HRIS) Specialist employed in a public higher education context, this researcher works with a set of legacy information systems that store backend human resource data including a state agency data warehouse, as well as several content management systems (CMS), information retrieval systems and a small number of disparate databases. Few of these information systems communicate with one another. Disparate systems may require businesses to undergo laborious data manipulation resulting in mistakes (Masayna, Koronios & Gao, 2009). As noted by McKnight (2009), non-integrated stand-alone systems can be problematic (p. 32) in that they present hidden costs such as (a) poor data quality due to redundant data storage (DAMA International, 2009; Marinos, 2004) with multiple versions of the same information causing confusion and inaccuracies (Marinos, 2004), and (b) mismatched or incompatible data configuration (Gibbs, Shanks & Lederman, 2005), for example differing business units capturing the same data elements in different formats (Buzydlowski, Hand, Song & Hassell, n.d.). According to Whitehead (2006), data quality is defined as "the degree to which data meets the following six key attributes: accuracy, reliability, credibility, timeliness, completeness and appropriateness" (p. 2). The desire to identify and avoid actions that lead to poor data quality and to identify processes which can be used to ensure and maintain high quality data is the impetus for this research study.

Businesses world-wide are struggling to keep pace with technology advances as the information growth rate exceeds storage capacity; setting record breaking growth by 62% in 2009 alone (Gantz & Reinsel, 2010). The digital universe is increasing at a phenomenal rate (DAMA International, 2009; Levi, 2008; Schutzer, 2010). According to Gantz and Reinsel

(2010) by 2020, the digital universe will increase 44 times the size it was in 2009 as institutions

continue to create and replicate unprecedented volumes of business intelligence (BI).This

presents a particular challenge for information management professionals who seek the best way

to manage the data assets of an organization. According to Whitehead (2006), data touches upon

all aspects an organization. Data has become "one of the most important" enterprise assets

(Jennings, 2004, p.61; Panian, 2010, p. 939; Redman, 2005, p. 3.), and yet corporations continue

to misunderstand or chose to accept enormous data inaccuracies (Olson, 2003; Whitehead,

2006). Many companies miss opportunities because they fail to recognize data as enterprise

assets (Badrakhan, 2010). As noted by Marco (2006a) "Most enterprises carefully manage other

assets (financial, physical and human) but overlook the immense value inherent in their data"

(p.28).

The need for quality data in the areas of revenue projection, cost reduction and compliance

assurance drives the demand for data governance processes (Panian, 2010). According to The

Data Management Association [DAMA] International (2009), there are data governance tools

that data management professionals can use to perform certain functions; "data governance tools

include data modeling, data management systems, data integration, quality tools, business

intelligence tools, document management tools, and meta-data repository tools" (p. 28). For the

purposes of this report, the term *data governance* is defined as "the formal orchestration of

people, processes, and technology to enable an organization to leverage data as an enterprise

asset" (Zornes, 2006, p. 1); and the terms *data* and *information* are used interchangeably in this

study. "Data quality is synonymous with information quality, since poor data quality results in

inaccurate information and poor business performance" (DAMA International, 2009, p. 291).

"Compliance as a data quality driver is more prominent in the post-Enron environment where state and federal regulations mandate strict data integrity" (Wolter, 2007).

Purpose

The purpose of this scholarly annotated bibliography is to identify and describe literature that provides ways to better manage enterprise-wide data assets within institutions of higher education through data quality actions and data governance options. The goal is to present a collage of selected data governance practices within the private business sector for consideration by individuals in public higher education who promote and support data quality initiatives. This study includes examination of literature as defined and applied in the private sector, based on (a) definitions of data quality concepts contributing to the need for data governance; (b) factors that drive data quality improvement principles and information governance; (c) descriptions of key data governance tools and practices, and (d) examples of data governance practices in select institutions of public higher education. Selected literature identifies actions or behaviors that impede the attainment and continuity of high quality data, defines the concept of data governance, and establishes what constitutes data governance best practices. The intent is to design the annotated bibliography in such a way that the audience can determine if a data governance program would improve data quality in public institutions of higher education.

Research Questions

Main research question. What can public institutions of higher education learn by examining approaches to data governance programs as defined and applied in the private sector?

Sub-questions.

1. What data quality conditions contribute to the need for a data governance program?

2. What factors drive data quality improvement principles and information governance?

3. How are data governance tools and practices defined?

4. How are data governance practices being used in select institutions of public higher education?

Significance

Enterprises rely on data with greater frequency for essential decision-making (Sarsfield, 2009), and strategic planning (Frost, Lucas & Blankert, 2004; Olson, 2003; Wende, 2007). Sarsfield (2009) states "success is increasingly tied to the quality of their information" (p.1). The cost of basing business decisions on questionable data can be enormous (Olson, 2003; Redman, 2005; Whitehead, 2006). Whitehead (2006) states that "data quality problems cost organizations millions of dollars, waste vast amounts of time and resources, and deceives management into making very poor decisions" (p.1).

Corporate scandals (McKnight, 2009) and problems concerning regulatory compliance focus greater attention on enterprise-wide data quality (Breur, 2009a; Marinos, 2004). According to Whitehead (2006), "many organizations acknowledge poor data quality is a major problem yet they accept it as inevitable" (p. 1). Institutions are riddled with data quality problems from inaccurate, missing, misinterpreted, and poorly defined information (Redman, 2005). As noted by Marinos (2004), it is important to understand the potential cost of poor data quality in order to mitigate risk. Data quality practices can help in this regard but usually lack the enterprise level oversight provided by a formal data governance program (McGilvray, 2006).

Although data governance is still an emerging concept (Badrakhan, 2010; Cheong & Chang, 2007), interest is picking up speed rapidly as information management professionals grapple with a monumental information growth rate (Gantz & Reinsel, 2010). According to Wende (2007) "both academic and practical sources presume data governance as a universal approach –

one that fits all enterprises alike" (p. 418). The same assumption is made for this study that data

governance may provide information quality solutions in institutions of public higher education.

Audience

The intended audience for this study is public higher education professionals in charge of

data asset oversight from a cross-organizational perspective. This can apply to information

management professionals in general, those new to the field, IT professionals collaborating with

business entities in a data quality or governance capacity, data quality practitioners, data

stewards, managers of institutional research in higher education, students, and executives

interested in the management of data as an enterprise asset. This study presents tools and

practical applications for the design, implementation and management of data governance

processes which can be used to ensure high-quality data assets.

Delimitations

Topic and focus. The goal is to select literature on data quality and governance that can be

described as general purpose in nature, rather than focusing on a specific industry. In addition,

literature is selected that examines data quality specifically in the public higher education sector.

Time-frame. Resources on data quality and governance span a period between 2003 and

2011, with few exceptions. References prior to 2003 are limited as the concept of data

governance is still fairly new (DAMA International, 2009), particularly in public higher

education administration.

Library access. This study incorporates reference material from a diverse selection of

electronic databases from multiple academic sources. Sources are accessed online through the

University of Oregon Library as well as the LaChance Library at Mount Wachusett Community

College. Key word searches are conducted on subjects and abstracts. Electronic documents are limited to full text in pdf, doc, rtf, and ppt formats.

Limitations

The range of content is restricted to topics that deal primarily with data quality and governance practices. Although there is reasonable amount of literature surrounding the topic of governance to ensure regulatory compliance, particularly in the health and financial sectors this paper does not fully examine the compliance facet, and only touches upon the topic. Therefore, a full discovery of this subject matter is outside the scope of this work. Data governance practices used in longitudinal data systems (LDS) is not covered in this study.

Reading and Organization Plan Preview

The reading plan arranges the selected literature into four primary categories which align with the overarching set of research questions. Categories include (a) definitions of data quality concepts contributing to the need for data governance; (b) factors that drive data quality improvement principles and information governance; (c) descriptions of key data governance tools and practices, and (d) examples of data governance practices in select institutions of public higher education. Abstracts and summaries are read initially as a way to determine the level of their appropriateness in relation to selection and evaluation criteria. Literature perceived to be a best fit is read in its entirety, guided by a set of detailed research questions (see Research Parameters).

Literature is organized in the Annotated Bibliography section of this paper first by research question category and then by author names. Each annotation includes a brief abstract, a more lengthy summary, and an assessment of the authority of the work.

Definitions

Definitions, according to Frost, Lucas, & Blankert (2004), play an integral role in accurately interpreting data. For this reason, the following section presents a list of specific definitions of terms as these are used in the literature selected for this study, in order to make clear their meaning.

Accuracy: "The closeness of measured values, observations or estimates to the real or true value" (Chapman, 2005, p. 3).

Basel II: Standard operating procedure originating in Europe which incorporates transparency of data handling in regard to key performance indicators (Sarsfield, 2009).

Business data stewards: "are recognized subject matter experts working with data management professionals on an ongoing basis to define and control data" (DAMA International, 2009, p. 40).

Business metadata: Data definitions which organize information into business subjects and describe individuals affected by the data (Jennings, 2004).

Coordinating data stewards: "lead and represent teams of business data stewards in discussions across teams and with executive data stewards" (DAMA International, 2009, p. 40).

Data: "The representation of facts as text, numbers, graphics, images, sound or video" (DAMA International, 2009, p. 2).

Data anomalies: "Missing data, near duplicates, and extraneous information all lead to additional costs when integrating data" (Sarsfield, 2009, p. 13).

Data asset: "In the information age, where data and information are the lifeblood of the 21st century, data is recognized as a vital enterprise asset" (DAMA International, 2009, p. 1).

Data decay: Data stays the same but over time accuracy of the data values decline (Olson, 2003).

Data enterprise asset: "Data to be shared and reused across multiple software applications and systems, business processes, and users throughout the organization" (Panian, 2010, p. 939).

Data governance: The DAMA International (2009) defines data governance as "the exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets" (p. 37).

Data governance maturity model: "Describes the journey from the AS IS to the SHOULD BE regarding the management of data, information and knowledge assets" (Sweden, 2009, p. 3).

Data lifecycle: The lifecycle of data is similar to other assets in that data moves through stages of storage, maintenance, use and destruction (DAMA International, 2009, p. 3; Khatri & Brown, 2010).

Data models: "specify the rules for any given database table" (Sarsfield, 2009, p. 25).

Data quality: "relevance, timeliness, completeness, trust and accessibility besides accuracy" (Breur, 2009a).

Data quality business rules: According to the DAMA International (2009), "data quality business rules is the process of instituting the measurement of conformance to specific business rules require definition" (p. 300).

Data quality management: "The collection, organization, storage, processing, and presentation of high-quality data" (Wende, 2007, p. 417).

Data steward: According to Villar (2009) the data steward is responsible for data definition including consistency, accuracy, and timeliness of critical information.

Data stewardship: "The formal accountability for business responsibilities ensuring effective control and use of data assets" (DAMA International, 2009, p. 39).

Decaying data quality: A gradual influx of errors made over a period of time that slowly erodes information quality (Breur, 2009b).

Disparate data (systems): Data elements spread out over multiple non-integrated systems (Sarsfield, 2009).

Executive data stewards: "senior managers who serve on a Data Governance Council" (DAMA International, 2009, p. 40).

Extract-transform-load (ETL): A tool to extract, transform and load key data elements from disparate sources into a single data warehouse (Petschulat, 2010; Sarsfield, 2009).

Fitness: "Refers to how suited data are for their intended use" (Breur, 2009a).

High quality data: "Data are of high quality if they are fit for their intended uses in operations, decision making and planning" (Lucas, 2010, p. 5).

IT assets: "Technologies that help support the automation of well-defined tasks" (Khatri, Carol & Brown, 2010, p. 148).

Information: "Information is data in context" of its current use (DAMA International, 2009, p. 2).

Information systems architecture: "Information systems architecture is ─the design of any complex technical object or system" (DAMA International, 2009, p. 65).

Information asset: (see Data asset).

Meta data: Data about the data inclusive of details to help interpret semantics of the data (Khatri & Brown, 2010).

Poor data quality: Inaccurate or misinterpreted data (Olson, 2003).

Stewardship: "A quality control discipline designed to ensure custodial care of data asset enhancement, risk management, and organizational control" (Sweden, 2009, p. 14).

System-Of-Record (SOR): The main system housing data considered to be the truth above all other systems (Breur, 2009b).

Research Parameters

The purpose of this section is to describe the research design used in the construction of this study. This section describes the context surrounding literature identification and categorization, key words used in the search strategy, and names of search engines and databases used in the research process. Evaluation criteria used to select literature are presented, along with a clearly defined reading and organization plan.

Search Report

The initial search of the literature includes collection of full-text articles on: (a) definitions of data quality concepts; (b) factors that drive data quality and improvement principles; (c) description of key data governance tools and practices; and (d) examples of data governance practices currently used in select institutions of public higher education. Sources searched for all search terms and phrases include academic journals, scholarly works, books, websites, professional white papers, and material from academic and professional meetings, conferences and reports. Sources include professional publications, case studies, white papers, reports, books, blogs, and articles. However, a gap exists in the literature pertaining to higher education institutions and the implementation of data governance practices.

This study uses the following research tools: Academic Index.net, Academic Search Premiere, Amazon.com, Business Search Premiere, Business Source Complete, Ebsco Host Database, ERIC, ERIC web portal, Gale Academic One File, Google, Google Books, Google Scholar, Google Wonder Wheel, Multiz Google, UniSA Research Archive, and professional websites; Ciber.com, Data Quality Campaign, DGS Club Express, Educause, Higher Education Funding Council (HERC), and Journal of Database Management in addition to select resources from reference material.

Documentation Approach

A list of collected references is categorized in tiers, I and II and entered alphabetically by author(s) in an Excel document. Tier I sources have the greatest relevance to the topic, such as peer reviewed works written by information management professionals, and professional white papers. Literature categorized as tier II is less relevant and its use reserved for supplemental background only. The process of reference collection seeks to identify:

1. Exemplar works, i.e., those publications cited repeatedly in the literature;

2. Works that target at least one element in my study;

3. Works available in full-text format.

Search Strategy

Key words. Initial search terms include: data governance, data quality and data quality in higher education. Key words derive from professional sources such as the DAMA International Guide, and from the keys words shown in the results of applicable findings.

- data quality

- data quality components and issues (private sector)

- data quality management

- data integrity

- data governance drivers

- data governance (framework)

- data governance roles and responsibilities

- data governance applications in higher education

- poor data quality

- quality records management

Initial search efforts utilize a number of different search tools: the University of Oregon (UO) Library and Catalogs, ERIC, Academic Index, BASE Academic Search Engine, EDUCASE, First Monday.org, Business Search Premiere, WhatIs.techtarget.com, Google Wonder World, Google Scholar, and Gale Academic One File. A preliminary review of the literature reveals it contains a substantial amount of current material on data quality and governance in respect to the financial and medical institutions. Less prevalent is literature on compliance and oversight processes in other areas, including public higher education institutions. As the search progresses and topic literature is found (or not) key categories transform. Data quality and data integration fold into data quality components and issues in the private sector; data quality management is now a separate category along with selected applications in higher education; data governance stewardship, tools, culture and enterprise perspective integrate with the data governance framework; and roles and responsibilities form a separate category.

Evaluation Criteria

In order to ensure credibility of resources the literature for this study is compiled from academic libraries, and professional websites and affiliates. Literature included in this report is scrutinized with regard to authority and objectivity of the source, quality of the material, coverage of the topic, and currency (Bell & Smith, 2007). Scholarly works contain proper citations and references; authors are examined for institutional affiliations and previous publications to determine professional competence; and the quality of publications produced by professional associations are considered for the value they bring to the community of specialization (Bell & Smith, 2007). Appendix A provides a breakdown of how each resource is categorized based upon the above stated evaluation criteria.

Reading and Organization Plan

This study is based on a detailed reading of a selected body of reference material related to data quality and governance. The reading plan is designed to meet both time and comprehension parameters. A preliminary reading of abstracts and summaries is conducted in order to determine level of appropriateness in relation to selection and evaluation criteria. Selected references are organized into four categories which correspond with the overarching set of research questions: (a) definitions of data quality concepts contributing to the need for data governance; (b) factors that drive data quality improvement principles and information governance; (c) descriptions of key data governance tools and practices; and (d) examples of data governance practices in select institutions of public higher education. Literature perceived to be a best fit is read in its entirety with the goal to address specific detailed questions, as described below. Through development of the research questions and sub-questions key words and concepts are identified which form the foundation for terms and concepts sought in the reference material, guided by a coding process described by Busch, De Maret, Flynn, Kellum, Le, Meyers, Saunders and White (2005). During the initial reading of each reference, this researcher determines the author's purpose, intended audience and main ideas; this information is catalogued in Excel format for ease of use and forms a base for development of the annotations.

The second read is conducted in a more detailed manner, designed to code key terms and concepts related to the core content areas. Similar phrases are analyzed to determine if they have the same meaning; those that do are combined into like categories according to a practice suggested by Busch et al (2005). Translation rules are established to ensure clarity of meaning, and discourage misinterpretation. Irrelevant information is discarded (Busch et al., 2005). As literature is thoroughly digested the physical material is pulled from the original alphabetizer and moved to a second alphabetizer for documents ready to be written about.

This second deep reading proceeds along the following plan.

Data quality concepts and relevance to data governance. When reading the selected literature addressing data quality conditions, this researcher seeks to answer the following questions:

What are the risks of not dealing with the issue of data quality?

What are the key benefits of data quality improvement and chief attributes of high data quality?

Data quality driving factors, improvement principles and information governance. When reading the selected literature addressing data quality improvement principles, this researcher seeks to answer the following questions:

How does continuous feedback factor into the data quality improvement process?

What is the relationship between IT governance and data governance?

What role does corporate maturity play in the decision to implement a data governance program?

Description of key data governance tools and practices. When reading the selected literature addressing data governance concepts, this researcher seeks to answer the following questions:

What are the common characteristics of selected data governance frameworks?

What are the reported data governance success factors?

Examples of data governance practices in selected institutions of public higher education. The last section of the Annotated Bibliography provides examples of data governance practices in selected institutions of public higher education. The goal is to find case

studies or literature reviews that deal with institutions of public higher education using data governance practices or tools to solve data quality problems.

Once the selected references are analyzed in relation to this plan, they are organized for presentation in the Annotated Bibliography section of this paper. The organization scheme utilizes the same four categories listed above in the reading plan, based on the research questions, as an underlying framework. The intent is to expound a set of best practices related to each of the four content areas by presenting a collection of credible sources sharing similar views on the same topic.

Data quality concepts. This category presents data quality as an enterprise asset upon which to capitalize. Aspects of best practice address:

- The root causes of data quality problems

- The importance of addressing the problems associated with poor data quality

- Characteristics of exemplary data quality

Driving factors and improvement principles. The content in this category presents compliance standards in the move toward data governance as an emerging trend in the private sector. Aspects of best practice address:

- The history behind the private sector trend toward data governance

- The impact of data governance on decision making to gain competitive advantage

Data governance tools and practices. This category presents the nuts and bolts of data governance as utilized in the private sector. Aspects of best practice address:

- Data governance as a business responsibility, and the importance of acting in close collaboration with IT

- Data as an asset to be leveraged across the entire organization

- The responsibility matrix, and the role of stewardship accountability on data quality

Data governance in public higher education. This category provides examples of how selected institutions of public higher education benefit from the implementation of data governance tools or practices. Aspects of best practice address:

- The need for data oversight in institutions of public higher education

- Data governance strategies to ensure institutional data quality

- The impact of data governance tools and practices on institutional strategic planning and decision-making

Annotated Bibliography

References in this section of the paper are organized into four categories. In the first category, *definitions of data quality concepts contributing to the need for data governance*, readers are informed of data quality problems which can occur in data warehouses and disparate silo systems often leading to adverse outcomes on information quality, and prompting the need for governance. The second category, *factors that drive data quality improvement principles and information governance*, brings the need for data governance into focus as it is implemented in the private sector, and identifies the subsequent impact on high data quality outcomes. The third category, *descriptions of key data governance tools and practices,* provides the audience with key methodologies of data governance to consider for possible implementation in institutions of public higher education. The fourth category, *examples of data governance practices in select institutions of public higher education*, recounts the experiences of several institutions of public higher education implementing data governance practices.

Annotations consist of three elements: (a) an excerpt from the publication; (b) an assessment of credibility; and (c) a summary of the content most relevant to this study. All ideas presented in the abstracts and summaries are credited to the author(s) of each reference. Paraphrased summary comments are not cited; direct quotes contain in-text citations indicating the page number of the reference.

Definitions of Data Quality Concepts Contributing to the Need for Data Governance as Applied and Defined in the Private Sector

Fisher, T. (2009). The data asset: How smart companies govern their data for business success. *John Wiley & Sons, Inc.,* Hoboken, N.J.

Abstract. Fisher posits that the notion that better data brings better decisions, leading to better business is easier said than done for most organizations that fail to implement a data governance program. Fisher explores how successful businesses are treating data and provides a guideline for building a data quality and governance business case along with methodologies on how to treat data strategically, as an asset. Fisher suggests data governance rules that can enhance data quality in addition to the introduction of a data governance program.

Summary. Despite billions of dollars spent on sophisticated information management technologies, Fisher states that corporate leaders are still being harmed by poorly managed, deficient or inaccessible data. "The quality, accessibility, and usability of data have an impact on every organization, but the issue rarely captures the attention of executives" (p. 5). Many corporations store data in disparate information systems across multiple departments which leads to a whole host of problems, and ultimately results in "a higher cost of doing business" (p. 65). Unmanaged data can lead to operational problems, poor decision-making, and reporting compliance issues. Even small data quality errors can have a huge impact on an organization's ability to maintain a competitive edge. Without quality data and integration, business operations will continue to be afflicted with data deficiencies, and the impact will span the entire organization. Data is considered an enterprise asset in proactive organization. Fisher identifies three core benefits to data quality improvement: (a) mitigating risk, (b) controlling costs, and (c) optimizing revenue. Well-managed information drives both productivity and innovation. Data

quality and governance must not be treated as a one-time project but as a continual process of oversight and assessment. Small improvements can make a large impact, according to Fisher who presents many examples of businesses transformations turning bad data into success stories by making nominal data quality improvements.

Credibility. Fisher is the president and CEO of DataFlux, a consulting agency helping companies to improve data, and establish and implement controls through integrated technologies. He speaks world-wide on data quality emerging trends, and information integration along with master data management and business optimization through better data management practices. This work contains a glossary of key terms, and well-cited references at the close of every chapter.

Lee, Y. W., Pipino, L. L., Funk, J. D., & Wang, R.Y. (2006). Journey to data quality.

Massachusetts Institute of Technology (MIT) Press, Cambridge, MA.

Abstract. All organizations deal with data quality problems from both systemic and

structural perspectives. The problem can't be solved with ad hoc methods or system level fixes.

This text provides a roadmap for executives and students to help guide them to plan and

implement a data quality management program.

Summary. All too often organizations seeking high-quality data develop new systems to

replace old ones but neglect to address the quality issues originally plaguing the old system.

Often, these ad hoc approaches lead to unsatisfactory results; rather than providing a solution, the

new system may exacerbate the problems of the old. It is critical for any organization to

communicate at all levels of the administration what high data quality is and what it can do: (a)

as a valuable asset, (b) to increase customer satisfaction, (c) to improve revenue, and (d) to

enhance strategic competitive advantage. There are ten root causes leading to data quality

conditions which include: 1) *multiple data sources*: producing different values for the same

information; 2) *subjective judgment*: wherein information is produced using subjective judgment

resulting in biased data; 3) *limited computing resources*: insufficient resources leading to

inaccessible resources; 4) *security/accessibility trade-off:* where access to information conflicts

with the policy set to protect it; 5) *coded data:* codes differ or conflict between various

disciplines making them difficult to decipher; 6) *complex data representations:* advanced

algorithms, in both image and text are inaccessible for content analysis in automated format; 7)

volume of data: magnitude of stored data creates challenge to access it in a timely manner; 8)

input rules: when input rules are restrictive data entry staff may arbitrarily change values to

bypass the control so they can enter the data. This type of inaccuracy is systemic in nature

making it more difficult to detect than a genuine error; 9) *changing data needs:* this type of

problem is two-fold, multiple consumers all with changing needs and all of which change

overtime with the potential to lead to deterioration of quality; and 10) *distributed heterogeneous*

systems: applications to integrate disparate systems may not have the proper integration

mechanism leading to data inconsistencies.

When data does not get used *believability* may be the issue. This can happen when the source

is not trusted due to poor reputation or the perception that the data adds little value. Permission

barriers and/or complex data representation, such as uninterruptable data codes, can lead to data

inaccessibility. Data is difficult to use if it is incomplete, inadequately defined, or inappropriately

integrated. Conventional controls alone are not enough to transform data quality problems; a

process-oriented technique is required in order to identify and correct where the operational

process has gone wrong.

Credibility. The authors are all data quality professionals with varying years of experience

contributing varying perspectives on data assessment approaches, quality policy setting,

challenges, and future trends. Lee is an Associate Professor with the College of Business

Administration at Northeastern University; Pipino holds a position at University of

Massachusetts as Professor Emeritus in the department of Management Information Systems;

Funk is the Founder of Beyond Accuracy, LLC and the Chief Information Architect; and Yang is

a professor at University of Arkansas, the Co-director of the Total Data Quality Management

Program at MIT, and Director of the MIT Information Quality Program.

Olson, J. (2003). *Data quality: The accuracy dimension*. San Francisco, CA: Morgan Kaufmann

Publishers.

Abstract. By the year 2020 electronic information is expected to grow to a staggering 35

trillion gigabytes as a result of multiple forms of media which joins the long list of data created

and replicated across the globe. Consequently, the workplace will feel the pain as information

management professionals struggle to store, secure and dispose of this large mass of electronic

content.

Summary. Data problems are magnified by replication and multi-system integration placing

a massive burden on content in the system of record (SOR). According to Olson, data is one of

the most important business assets. Poor data quality costs businesses anywhere from 15-25% of

profit. Poor data management is said to cost worldwide business $1.4 billion each year in

corrections to billing, accounting and inventory. A large portion of that cost is due to data quality

inaccuracies. The following information characterizes the typical organization's awareness and

responsiveness to data quality:

They are aware of data problems.

They consistently underestimate, on a considerable scale, the extent of the problems.

They are unaware of the cost to the organization in relation to the problems.

They are unaware of the potential value to be gained from fixing the problems.

They typically blame the IT department for poor data quality, even though a good deal of the

problem stems from outside the IT department in the way of poorly articulated requirements,

tolerance of testing systems, and poor data generation processes, to name a few. As noted, "Data

quality problems are universal in just about any large organization. The fact that data quality is

universally poor indicates that it is not the fault of individually poorly managed organizations,

but rather that it is the natural result of the evolution of information system technology." (p. 10).

Two key factors contribute to poor data quality: 1) rapid system implementations which can

challenge quality control efforts; and 2) slowly evolving quality control tools that don't keep

pace with systems they support. Correction activities are a normal part of doing business and not

viewed as a *data quality issue*. The activity grows and without much fanfare; and since staff are

not working in isolation doing just that one job the scope of the problem is usually not

recognized. Also, corporate leaders want to believe they have top notch information systems;

"they do not want to expose to their board or to the outside world the facts of inefficiencies or

lost opportunities caused by inaccurate data" (p. 11). The impact of poor data quality includes

transactions costs of rework, cost to implement new systems, delays in conveying information to

decision makers, lost customers due to poor service, and lost production. Characteristics of data

quality include accuracy, timeliness, and relevance, in addition to making sure the data is

complete, understood, and trusted. Olson suggests that every organization needs a methodology

to regularly monitor and improve data quality from within their information systems.

Credibility. Olson has an undergraduate degree from Illinois Institute of Technology and an

MBA from Northwestern University. He has a 36-year history in data management systems and

is widely recognized as an expert in database technology. In addition to this publication, Olson

published a book in 2008 titled Database Archiving: How to Keep Lots of Data for a Very Long

Time. In 2003 Olson was the chief technology officer and VP of engineering at Evoke Software.

Redman, T. C. (2001). Data quality: The field guide. *Butterworth-Heinemann*, Woburn, MA

01801-2041.

Abstract. The book provides guidance on data quality program implementation. Coverage

includes data quality problems and improvement processes. The text incorporates what those

institutions "with the best data" do and examines social barriers to successful data handling

practices.

Summary. Every business generates enormous quantities of data; when wrong, bad data

costs these organizations in time, revenue and reputation. Based on a handful of carefully

conducted proprietary studies, Redman suggests 10% of revenue is impacted by poor data

quality; a figure that doesn't even include the cost of bad decisions or depleting morale which is

harder to measure yet of greater importance. Those at the top of the administration need to take

the lead. "CEOs thought they could establish broad goals, make the right speeches, and leave

everything else to their subordinates… they didn't realize that fixing quality meant fixing whole

companies, a task that can't be delegated" (Juran, 1993, p. 7). In chapter 7.1 Redman leads with

the tag line "Take this personally" in regard to the dangers of liability that manager's face when

failing to provide quality data, or simply neglecting to make good on regulatory compliance.

Under these circumstances the objective of the data quality initiative should be to avert

embarrassment. Additionally, poor data quality impacts decision making which, over time, can

lead to a lack of trust in the data making it more challenging to align the organization behind

business decisions. To improve data quality you must find and fix errors, and examine and

implement ways to prevent them by eliminating the source of corruption.

Credibility. Redman is a trained statistician starting his career over thirty years ago working

for Bell Labs where he focused his attention on data quality, a good ten years before anyone

acknowledged there was an issue. In 1996 he established his own business, Navesink Consulting

Group, assisting organizations in a data management capacity as an experienced data quality

consultant. Redman is quoted extensively in data quality circles. He has a PhD in statistics from

Florida State University.

Redman, T.C. (2008). Data driven: Profiting from your most important business asset.

Retrieved from: *Harvard Business School Publishing*, Boston, MA 02163.

Abstract. This book targets business managers and is broken into four sections; (a) simple

steps to get started, (b) how to craft a midterm plan, (c) plan to think long term, and (d) changing

the culture by getting buy-in. This text outlines the ways in which companies can optimize

business practices to achieve greater profit from data--what Redman coins throughout the text as

"your most important business asset". The material presents a roadmap on how businesses can

get their organizations to recognize the importance of data quality.

Summary. According to Redman, "bad data land organizations in deep trouble with

disturbing regularity" (p. 35). Organizations must take as much care with their data as they do

with their capitol and human assets, "which for data and information better care is mostly about

quality" (p. 3). Redman identifies the seven most common data quality issues as:

1) *People can't find the data they need:* People are frequently unable to locate the data they

need according to Redman, "wasting 15-30% of their time in the search process, with a 50%

or less success rate" (Feldman, 2004);

2) *Incorrect data:* Another sobering statistic is that incorrect data plagues businesses10-25%

of the time, according to a Gartner study (as cited in Whiting 2006);

3) *Poor data definition:* Data is repeatedly misinterpreted with no ability to make

connections between departments;

4) *Data privacy/data security:* Data has many attributes including that employees can easily

share it and transport it in electronic format at a low cost. This ease of use presents problems

in that it can be vulnerable to theft. The data can then be stolen without anyone knowing--by

simply making an electronic copy. Under more non-digital circumstances you would be

aware when an asset was stolen—not necessarily the case when the data is electronic;

5) *Data inconsistency across sources:* Data inconsistency stems from the same data being

stored in multiple locations in different formats;

6) *Too much data:* Uncontrolled data redundancy and collecting that is never used leads are

just several examples of too much data; and

7) *Organizational confusion:* Many organizations are not able to answer the simplest

questions about their data such as *where is all the data stored* and *what is the value*?

Credibility. Redman is a trained statistician starting his career over thirty years ago working

for Bell Labs where he focused his attention on data quality, a good ten years before anyone

acknowledged data quality was an issue. In 1996 he established his own business, Navesink

Consulting Group, assisting organizations in a data management capacity as an experienced data

quality consultant. Redman is quoted extensively in data quality circles. He has a PhD in

statistics from Florida State University.

Sarsfield, S. (2009). *The data governance imperative. A business strategy for corporate data.*

Cambridgeshire, United Kingdom: IT Governance Publishing.

Abstract. Based on the premise that the ability to produce accurate enterprise information is

more important than ever, this book is written from the perspective of a business person's

examination of data governance and includes strategies and tactics important to data champions.

Businesses rely on better operational processes, and a commitment to high-quality, clean data.

Both internal and external data can be a potential opportunity to enhance business intelligence,

speed and agility beyond that of the competition.

Summary. Sarsfield examines the need for data governance. Corporate success is

increasingly linked to data quality and the ability to make reliable intelligent business decisions.

Companies evolve information handling processes over time; in the early stages, "everyone

makes their own rules and owns their own data. In stage one there is no cohesive plan for data,

nor is there anyone particularly responsible for managing data" (p. 1). As the company matures,

the amount of information escalates, and the need for effective information handling practices

increases. Companies don't tend to consider the information challenges of corporate growth;

"it's usually cheaper and easier to fix an immediate problem with glue and duct tape than it is to

think about it strategically" (p. 3). As companies undergo the data maturation process, they begin

to realize certain inefficiencies: (a) data silos form with each business unit acting an island

developing individual data handling strategies; and the data governance maturation process

undergoes a certain metamorphosis; (b) data anomalies begin to proliferate; (c) the

implementation of ad hoc solutions perpetuate problems since they address a specific business

unit's issue and not those of the entire enterprise. Sarsfield presents the following factors that

hamper the successful acquisition of business intelligence: lack of standards, typos and

duplications, multiple platforms, varying languages, internal competition to one-up another business unit with stellar information quality processes, data age, information reliability, and unknown data.

Credibility. Sarsfield works in product marketing where he specializes in strategic planning and development at Talend, a company that produces open source software solutions to businesses with an interest in data integration and performance excellence. His professional experience began in 1987 as the Managing Editor of Mindcraft Publishing. Sarsfield has spent the past thirteen years, in various capacities, working in the computer software industry.

Weller, A. (2008). Data governance: Supporting datacentric risk management. *Journal of*

Securities Operations & Custody, 1(3). Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=339

38168&site=ehost-live&scope=site

Abstract. This paper addresses trends in data usage and the implications for how business is

likely to be impacted in the years forthcoming. The work covers key considerations for data

governance program implementation. This article may be of particular interest to readers new to

the topic of data governance or those who are familiar with the concept but seeking more

program implementation support.

Summary. Accurate information is crucial to corporations in the financial sector. Problems

arise and are compounded by things like data duplication between information systems. Many

institutions have no idea of the enormity of data they generate and are ignorant of data

management methods being used that no longer work in the information age. "In addition to

sheer volume of data produced, intranets, wikis and other collaboration technologies have driven

information out of filing cabinets and online, making vast amount of data accessible anytime,

anywhere" (p. 251).

At the same time, storage techniques are proving substantially less durable than their paper

storage predecessors. Data is constantly moving between organizations in today's environment,

and across multiple systems lacking a consistent margin of protection. Although there are

significant benefits to electronic information storage there are substantial new risks. The impact

on decision-making could have catastrophic outcomes. In order to maximize the benefits of

information organizations should recognize, and act on trends suggesting increased risk; this is

the goal of data governance. Core benefits from the implementation of data governance practices

could include: (a) the acquisition and maintenance of high-quality data with which to make decisions; (b) the ability to successfully manage large volumes of stored and newly created data; (c) achievement of an understanding of how information flows within the organization and potential sources of overlaps in data stores leading to an opportunity to investigate consolidation strategies; (d) obtaining the utmost value from data assets; (e) ensuring privacy and security of sensitive information; (f) meeting compliance requirements to limit fines and reduce the potential damage to the organization's reputation; and (g) reducing costs associated with need to generate evidence for litigation.

Credibility. Weller is employed by Protiviti, an international business consulting firm specializing in technology, risk, compliance and governance issues, where he is in charge of Security and Privacy for the Pacific Northwest, and is a leadership member of their global security team. He has regular speaking engagements at conferences, and has written articles published in several trade publications. Weller has a BS from the University of Manchester England, and is currently working on a certificate from the University of Washington in Information Assurance and Cybersecurity.

Factors that Drive Data Quality, Improvement Principles and Information Governance as Defined and Applied in the Private Sector

Breur, T. (2009a). Data quality is everyone's business--designing quality into your data

warehouse--Part 1. *Journal of Direct, Data and Digital Marketing Practice, 11* (1)

doi:10.1057/dddmp.2009.14. Retrieved from

http://find.galegroup.com.ezmw.ez.cwmars.org:4200/gtx/infomark.do?&contentSet=IAC-

Documents&type=retrieve&tabID=T002&prodId=AONE&docId=A205907920&source=gal

e&srcprod=AONE&userGroupName=mlin_c_wachcc&version=1.0

Abstract. Data quality, despite monumental volumes being generated in the information age,

can still produce a competitive advantage for those with the foresight to address the challenges

with creative data management practices. Combining data from silo systems presents an

opportunity to create new data streams from which valuable business intelligence can be mined.

However, the process of system integration has a tendency to reveal the previously unchecked

proliferation of data quality issues. This work proposes a data modeling solution (Data Vault)

and development methodology (Agile), to provide deal with these problems.

Summary. Breur begins this article by addressing the evolution of new reporting

complexities; a phenomenon resulting from the introduction of the Sarbanes-Oxley Act in 2002,

in conjunction with the need to digest information in a timely manner. Breur further attributes

new-found interest in data quality on the need to stay competitive, and the desire to capitalize on

data asset opportunities. To address the impact of production data, Breur points to three

elements: (a) the tremendous increasing volume of data, and new data sources persistently being

developed, (b) *change* as it applies to the continual transformation of business intelligence

solutions, and (c) dissatisfaction in return on investment (ROI) from information technology

investments. Since a great deal of data is available publically, companies are finding it difficult

to maintain a competitive advantage. With so much information out on the internet for anyone to

digest, in addition to other publically accessible information, organizations quickly learn of

competitors new innovations and work relentlessly to one up them. There is one source of data

that should not fall prey to the competition, proprietary data. For this reason, companies should

extract their proprietary data with care, ensure high quality, and protect it as an enterprise asset in

order to maximize value "as a sustainable competitive advantage" (p. 21).

Credibility. Breur is the principal of XLNT Consulting. He is an expert in the business

aspect of data mining and speaks regularly at workshops and international conferences. Breur has

spent the past ten years specializing in advising businesses on how to utilize their data to achieve

the best possible outcome. He teaches an MBA Program for Certified Business Intelligence

Professional (CBIP) at universities. Breur is affiliated with numerous data mining and financial

institutions and was cited by Harvard Management Update for his state-of-the-art data analytics.

Breur, T. (2009b). Data quality is everyone's business – managing information quality—Part 2.

Journal of Direct, Data and Digital Marketing Practice, 11 (2) 114-123. Retrieved from

http://www.palgrave-journals.com/dddmp/journal/v11/n2/pdf/dddmp200921a.pdf

Abstract. This article examines maintenance and ongoing measures that ensure data quality.

Breur distinguishes between a data quality project wherein a short-term operation is required to

address non-quality problems and an ongoing data quality program which is more appropriate to

address the *root* of non-quality from an operational perspective. Both issues often occur

simultaneously and therefore joint approaches are called for. Ideally, as the data lifecycle runs its

course, sources of non-quality are identified, associated organizational costs are uncovered,

training is implemented, and awareness spreads across the organization while tools and

technology to support the process align with accountabilities to produce quality.

Summary. According to Breur, in order to diminish the likelihood of enabling poor data

quality and consequent expense, and to ensure the generation of value from data, a concerted

effort is required across the entire organization. The implementation of a business rule will

enable staff to make the systemic decision about which data format to follow by mirroring the

format contained in a System-Of-Record SOR. The lifecycle or maturity of a company has a

direct relationship to how the business addresses the issue of data governance. In the early stages

a business is more apt to welcome more creativity and less restriction in order to promote

innovation and jump on advancing opportunities. Freedom to act quickly making decisions in

such an environment tends to trump cost controls. As a company ages, IT expenses grow, and the

need to control costs becomes much more apparent. Breur speaks on the difference between one-

time data projects and ongoing data quality programs. The distinction is made when data quality

projects are designed to correct a mass of information within a certain period of time vs. a data

quality program which is a continual process involving oversight and assessment. Breur provides

the example of data irritants that waste time and build up gradually to be quite costly. An

example might be when staff members repeatedly have to look things up in external systems

because the information isn't retained locally, or when incorrect data, such as a phone number,

causes a failed attempt at making contact with someone. Breuer recounts his success with the

data quality scorecard, a tool designed to support data quality programs. The scorecard provides

statistically stratified samples to keep track of erroneous duplicate entries. The last objective,

according to Breur, is the need to align business goals to make data quality the norm. The author

uses the example of ending the practice of rewarding employees for speed, not accuracy.

Credibility. Breur is the principal of XLNT Consulting. He is an expert in the business

aspect of data mining and speaks regularly at workshops and international conferences. Breur has

spent the past ten years specializing in advising businesses on how to utilize their data to achieve

the best possible outcome. He teaches an MBA Program for Certified Business Intelligence

Professional (CBIP) at universities. Breur is affiliated with numerous data mining and financial

institutions and was cited by Harvard Management Update for his state-of-the-art data analytics.

Lucas, A. (2010). Corporate data quality management in context. *LaboratorioNacional de*

Engenharia Civil and Instituto Superior de Ecohomia e Gastaok Portugal (research in

progress). Retrieved from http://hdl.handle.net/10400.5/2271.

Abstract. This paper examines the strategies organizations use to deal with data quality

management. It includes case-based information on data quality cost factors, how business

drivers motivate stakeholders to take on data quality initiatives, and various concepts, roles and

responsibilities in the data quality management approach.

Summary. Problems with data quality and management practices are of growing concern to

both the academic and professional communities. Poor data quality costs can be broken into

three categories: 1) *process failure costs:* this happens when operational processes are of such

poor quality that outcomes fail to meet business objectives as in the example of an inaccurate

shipping label causing a product to be mis-delivered; 2) *Information scrap or rework*: such as

business costs that occur when an item is re-mailed, or discarding defective data to produce a

quality outcome; and 3) *opportunity costs* as a result of lost revenue. For example, low accuracy

of the address data may lead to the misdirection of customer loyalty-cards where a certain

percentage will never receive fundraising or advertising campaigns. "To be effective, data

quality management must go beyond the activities of *fixing non-quality* data, to *preventing data*

quality problems by managing data over its lifecycle to meet the information needs of their

stakeholders" (p. 2). Assessment and improvement are two key common phrases that lead to

diagnosis of quality data, and define relevant quality dimensions. Improvement primarily entails:

(a) identification of underlying causes that cause poor quality data to proliferate; (b) error

correction using data quality tools; and (c) the redesign of processes that generate or modify

information to improve quality.

Credibility. Lucas is a Visiting Associate Professor at Instituto Superior de Economia e Gestão in Lisboa, where she is preparing her PhD in Management. She holds a masters degree in Computer Science from Université Scientifique et Medicale de Grenobolekl France. Lucas has published professional articles, reports, books and videos spanning 1973 to 2010.

Masayna, V., Koronios, A., & Gao, J. (2009). A framework for the development of the

business case for the introduction of data quality program linked to corporate KPIs &

governance. *2009 Fourth International Conference on Cooperation and Promotion of*

Information Resources in Science and Technology, pp.230-235. Retrieved from

http://doi.ieeecomputersociety.org/10.1109/COINFO.2009.56.

Abstract. Measuring efficiency and effectiveness of processes is done using key

performance indicators (KPIs). KPIs are important to decision-making as management relies on

these measures, at all levels of the enterprise, to attain successful outcomes. Organizations need

the ability to identify high-worth, high-risk issues with data quality as they support each KPI.

Thus, a rigorous guiding structure that maps data quality aspects to KPS will promote decision-

making confidence. Such a structure is additionally critical in assisting businesses to develop and

improve data quality objectives at multiple business levels.

Summary. This paper presents the results of a national survey and interview process

conducted to identify a framework for connecting data quality endeavors to organizational KPI's.

KPI's can be misleading and likely to produce erroneous outcomes if the indicator is based on

poor data quality. Inadequate data processing may lead to KPI-related problems. "For example,

disparate or interfaced software systems from different manufacturers may require organizations

to undertake laborious data manipulation to compile KPIs; data transfer between different

software systems could result in mistakes and inaccurate KPI calculations" (Hoover & Schubert,

n.d., p. 2). In addition, data is often held in non-integrated and unrelated information systems in

inconsistent formats (Neely & Bourne, 2000). These findings identify a solid correlation between

performance measurement, KPIs and data quality initiatives. The key to performance

improvement is effective data quality management. However, this study indicates a lack of

methodical guidelines to assist organizations in building robust data quality proposals. Resource

limitations and complex processes make it difficult to develop, execute and prioritize sufficient

data quality programs. The following framework (figure 1) illustrates the link between data

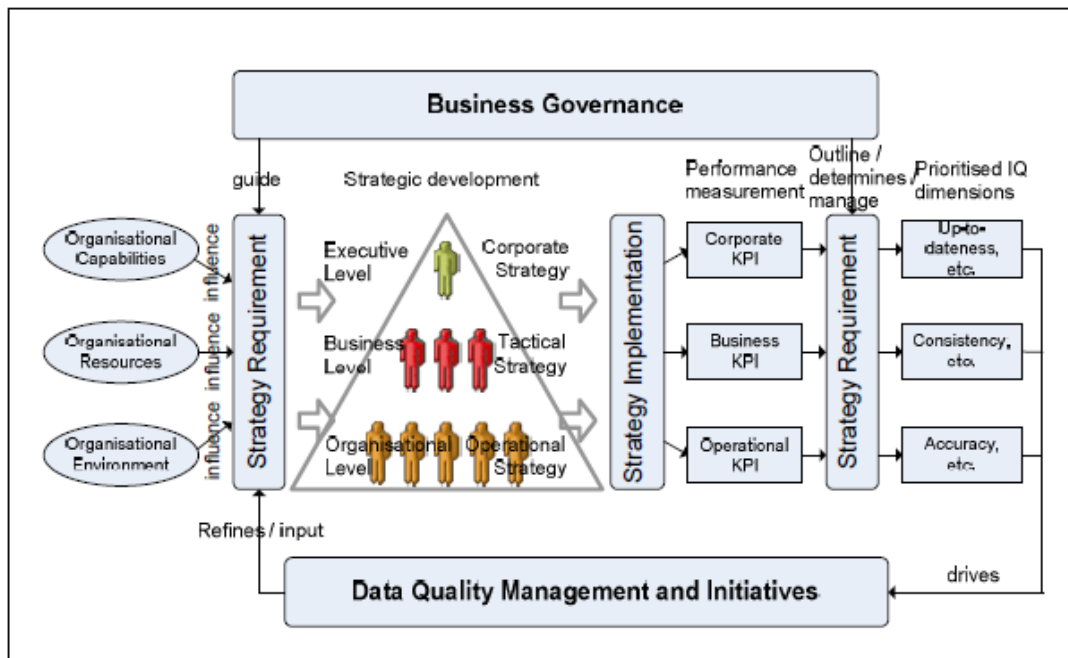quality proposals, governance and KPI's as derived from survey analytics and interviews.

Figure 1. Linking corporate KPIs, governance and data quality initiatives (p. 234).

Credibility. Masayna is a lecturer with the Department of Business Computing Faculty of

Accountancy and Management at Mahasarakham University in Thailand, holding a PhD from

the University of South Australia. Masayna is the recipient of professional awards, has teaching

experience, and is affiliated with relevant boards and committees. Koronios works for the School

of Computer and Information Science as the Head of the Division of Information Technology,

Engineering and the Environment. Information quality and governance is one of his professional

interests. He holds a PhD from the University of Queensland. Gao is a Senior Lecturer at the same institution and division as Koronios. He is published in professional journals and conference papers on topics specifically related to data quality.

Otto, B., Wende, K., Schmidt, A., & Osl, P. (2007, Dec). *Towards a Framework for Corporate*

Data Quality Management. Paper presented at the 18th Australasian Conference on

Information System, Toowoomba, Australia. Retrieved from

http://aisel.aisnet.org/acis2007/109

Abstract. This paper suggests a framework for managing corporate data quality comprised

of three areas from the perspectives of data governance and execution: strategy, organization and

information systems. The structure helps readers determine which tasks will improve enterprise

data quality and understand how governance and execution are interrelated. Business objectives

in combination with corporate data management provide an anchor to the current organizational

structure.

  Summary. Companies need to acclimatize to new business models in order to attain

operational excellence from a process efficiency and business effectiveness perspective. This

framework contains two perspectives, one of governance and the other of execution (figure 2).

Governance refers to what needs to be accomplished, who needs to be involved and what

responsibilities are assigned and distributed. Elements of the governance structure are strategic,

organizational and system-related and shape the decision-making underpinning for the entire

corporate data quality program. Execution is the point where specific functions and activities are

implemented such as follow-through on data quality policies, monitoring data quality and

problem solving. Therefore, governance and execution are inter-related forming a control loop

for continuously reviewing and adjusting elements as necessary to ensure the high quality data.

  "All data quality policies, practices, principles, standards and the data quality architecture

reflect the business view" (p. 921). Corporate data quality governance practices contain three key

layers, strategy, organization and architecture. A data quality strategy is required to manage all

activities related to data quality maintaining alignment with the global business strategy. The

strategy incorporates business goals, analyzes stakeholder functions and the role of data in the

organization. In addition, the strategy defines the portfolio of data quality initiatives. Design of

the corporate data quality organization entails the determination of consumer information

requirements, defining data manufacturing procedures, defining roles and responsibilities over

divisional boundaries, specifying metrics and standards for data quality, and establishing policies

and procedures. The corporate data quality architecture involves developing a common data

object model, creating a data dictionary from the business perspective, and defining support for

information systems.

Figure 2. Illustrates six practices in the CDQ Framework (p. 920).

Credibility. Dr. Otto is an assistant Professor at the University of St.Gallen's Institute of

Information Management of Corporate Data Quality. He has published in scientific journals,

conference proceedings, books, and delivered invited presentations. Wende holds a doctorate of

information management through the University of St. Gallen. Her primary professional focus is

on data quality management through corporate data governance. Schmidt is a PhD student

working for the Hasso-Plattner-Institute (HPI). His main field is research is in monitoring

operating systems to support fault-tolerant distributed applications. In addition he created an

inspection framework which consistently accesses data structures within operating systems. Dr.

Osl is from the Institute of Information Management in Gallen, Switzerland. His university status

is listed as post-doctor and he has published journal papers, book chapters, conference papers,

and case studies.

Panian, Z. (2010). Some practical experiences in data governance. *World Academy of Science Engineering and Technology.* Summary retrieved from
http://www.waset.ac.nz/journals/waset/v62/v62-174.pdf

Abstract. Businesses, small, medium and large are resolute in their effort to manage data as valuable enterprise assets, shared with users across the entire organization from multiple sources and systems. Practitioners are working to determine standards, and establish policies and processes for effective information management. These professionals recognize the critical need to develop the right foundation to support an effective data governance program. This article focuses on data governance initiatives and conventional business drivers.

Summary. Panian describes data as "one of the most important assets in an organization" (p. 939). This work ties data governance initiatives with the common business drivers that support them, and describes the relationship between business and IT. There are six key attributes enterprise-wide data must contain: accessibility, availability, quality, consistency, auditability, and security. Data governance expectations include (a) the need to ensure that business needs are met, (b) to protect data and manage it as a valuable enterprise asset, and (c) to reduce the costs of managing data. Business drivers of data governance initiatives include: (a) increasing revenue growth rates and retention numbers, (b) reducing costs by improving operational efficiency through automation of business processes and elimination of redundancy, and (c) ensuring regulatory compliance through internal governance policies, and by streamlining reporting and auditing processes. Data governance provides the tools to improve data quality through oversight practices and technology innovation. The data integration lifecycle provides the opportunity to identify new quality issues for resolution and ensures accountability for critical data metrics. Successful data governance initiatives have a strong working relationship between business and

IT. "In most cases, the business assumes ownership of the data taking the lead in driving data governance; an appropriate relationship, since ultimately the data exist to serve the business and the business is the primary beneficiary of effective data governance" (p. 944). IT then collaborates with business to introduce the most appropriate technology in support of the data governance program. Together, IT and business pool resources to establish and track specific metrics which must demonstrate clear business benefits to support the program. It is at this juncture that visibility is heightened and program awareness begins to surface throughout the organization.

    Credibility. Panian is a professor of informatics at the Faculty and Economics and Business, University of Zagreb in Croatia where he also holds a PhD from the same institution. His professional interests focus on enterprise information systems in addition to e-business and business intelligence. His publications include 32 books and 150 professional papers. Panian is also an international keynote speaker.

Power, M. E., & Trope, R. L. (2005). Sailing in dangerous waters: A director's guide to data

governance. *American Bar Association, Chicago, IL 60610.*

Abstract. Corporations are faced with ever changing data-related obligations. Legal drivers

and threats are examined in this work in relation to management's data governance efforts as

they apply to information asset security. The reference provides a guide for corporate directors

and other leaders to help them fulfill their personal and organizational obligations in regard to

data governance.

Summary. The digital era is the driving force behind the mandate to regulate corporate

information causing heightened concern in the form of new federal, state and international

regulations on data governance practices, especially as they relate to information security. The

Sarbanes-Oxley Act 2002 (SOX), in particular, set new standards for public companies and their

internal accounting controls. SOX is significant from a data governance perspective with its

emphasis on the "creation, evaluation, assessment and correction of… internal controls for

financial reporting" (p. 18).

Directors and those responsible for the security of corporate data risk personal liability if they

do not ensure the integrity of their organization's data. Damage control can be expensive and the

cost to a company's reputation can be substantial. A cross-organizational approach to data

governance should be encouraged making data integrity and security a vision shared across the

entire organization. Executives and leaders that extend data governance beyond the parameters

enforced by regulatory compliance will be able to gain a competitive edge over those lagging

behind or those restricting oversight actions to the bare minimum.

Credibility. Power is a Chief Privacy Officer and experienced partner and legal advisor with

the law offices of Gowling, Lafleur, and Henderson LLP. Trope is an Adjunct Professor at West

Point in the Department of Law, and co-editor of the Digital Protection Department and Privacy

Department, journal IEEE Security & Privacy, and serves as an advisor on data governance

management and information security practices. With a JD from Yale Law School, Trope

coauthored *Checkpoints in Cyberspace: Best Practices for averting Liability*, 2005. Both Power

and Trope have professional affiliation; Power, is a member of the National Executive of the

Privacy Law Section of the Canadian Bar Association, the Canadian Information Technology

Law Association, and the American Bar Association's Cyberspace Law Committee. Trope is a

member of the Association of the Bar of the City of New York's Information Technology

Committee and the American Bar Association's Cyberspace Law Committee. The authors

provide relevant properly cited resources.

Wallace, L., Lin, H., & Cefaratti, M. (2011). Information security and Sarbanes-Oxley

compliance: An exploratory study. *Journal of Information Systems*, *25* (1), 185-211. Abstract

retrieved from

http://content.epnet.com/pdf25_26/pdf/2011/JIN/01Mar11/59414945.pdf?T=P&P=AN&K=5

9414945&EbscoContent=dGJyMNLr40SeprM4v%2BbwOLCmr0meqK5Ssaa4Sq%2BWxW

XS&ContentCustomer=dGJyMPGuslGwrbVNuePfgeyx%2BEu3q64A&D=aph

Abstract. This paper focuses on the resurgence of interest in information controls as a result

of the Sarbanes-Oxley Act of 2002 (SOX). The study is based on survey data collected from 636

participants from the Institute of Internal Auditors (IIA). It explores the extent that IT ISO 17799

security controls are embedded into the administration's internal control environments.

Summary. Congress is responsible for enacting SOX to protect shareholders as well as the

public from deceptive enterprise practices and financial errors. The accuracy of financial

information and success of internal controls are the primary focus of SOX. As a result of SOX

compliance companies are increasing transparency, reliability and accountability of financial

information.

"Another control is ISO 17799, the International Standard for the Code of Practice for

Information Security Management, which provides a detailed list of controls that can be used for

establishing an information security program" (p. 186). ISO 17799 provides formal guidance on

how to integrate IT with SOX initiatives. Many organizations adopt an ISO 17799 framework in

their information security and maintenance plans. Sections 3 through 12 of the ISO standard

address ten control categories: 1) *security policy* in category 3 covers the creation and

implementation of security policies; 2) *organizational security* in category 4 covers the

development of an enterprise-wide information security infrastructure; 3) *asset control and*

classification in category 5 is meant to protect enterprise assets, "including maintaining an

inventory of organizational assets" (p. 189); 4) category 6 deals with *personal security*. This

classification deals with reducing the potential for loss of information by way of human error,

misuse or fraud; 5) *physical and environmental security* in category 7 addresses protection of

critical information assets as well as other physical assets; 6) *operations management and

communications* in 8 concentrates on the implementation of operational controls; 7) category 9,

access control, addresses protecting access to an organization's information; 8) *systems

development and maintenance* in category 10 deals with building information security processes

into information systems; 9) *business continuity management* from category 11 involves disaster

recovery planning for minimal interruption should a disaster strike; and 10) category 11

addresses *compliance* to avoid breach of criminal or civil regulations.

Credibility. The article is published in a peer-reviewed journal. Wallace is an Associate

Professor at Virginia Tech in the Department of Accounting and Information Systems. She has a

PhD from George State University in Computer Information Systems. Wallace is published in

Decision Sciences, Communications of the ACM, Information and Management, IEEE Security

& Privacy, and Journal of Systems and Software. Lin has a PhD in Accounting and Information

Systems from Pamplin College of Business. He is an Assistant Professor from DePaul University

where he works in the School of Accountancy and MIS. Lin's research is in knowledge

management, IT governance and internal control. Cefaratti is a faculty member at Northern

Illinois University where she teaches financial accounting and assurance services. She is the

recipient of awards from the Accounting and Information Systems Educators Association.

Description of Key Data Governance Tools and Practices as Applied and Defined in the

Private Sector

Berson, A., & Dubov, L. (2011). Master data management and data governance. *McGraw-Hill*

Osborne Media; 2 editions.

Abstract. This book examines current data architecture and technology viewpoints along with system development and information management methodologies. Instruction on designing a business case around the need for Master Data Management (MDM) is presented. In addition, information is provided on building accurate data models, executing layered security policies and challenges that go with the integration of legacy data systems.

Summary. Data governance as a discipline suffers from two widely held negative beliefs: (a) that data governance is simply a buzz-word for other somewhat similar applications such as business process improvement, business analysis, and data quality improvement, and (b) if left uncontrolled data governance programs can become too large, bureaucratic, and inefficient. However, these beliefs about data governance are too vague, and issues of bureaucracy are addressed when administrators take a best practices approach to data governance. Examples of effective methods of data governance might include: (a) the establishment of data councils and boards holding regularly scheduled meetings, and (b) institute communities of practice with a focus on data quality, metadata, data modeling, and data protection, all under the header of data governance. A risk-based objective, for example, will cultivate early detection and lessen the likelihood of costly issues in terms of threats and opportunities that could materialize over time.

Most companies report some degree of data governance, primarily around data security and in support of compliance requirements such as Basel II. However, "Most organizations do not see a way or don't feel they have a need to reach higher levels of data governance maturity...

[which is] consistent with what we see in the field" (p. 400). A data governance framework is required in order to break large-scale, cross functional, cross-organizational, and cross-systems programs into structured, manageable chunks.

There are many different frameworks to choose from; several are described as follows: (a) Mike 2.0 Framework: an open-source methodology with a focus on information development. It includes a data governance strategy, organization, policies, processes, investigation and monitoring, and technology and architecture; (b) the Data Governance Institute Framework covers policies, strategies and standards with a primary focus in Business Process Reengineering (BPR), and Enterprise Data Management (EDM), data quality, privacy, compliance and security, integration architecture, business intelligence, and data warehousing; and (c) the IBM Data Governance Council Framework and Maturity Model with a framework that consists of organizational structure and awareness, stewardship, policy, value generation, data risk management and compliance, information security and privacy, data architecture, data quality management, classifications and metadata, information lifecycle management, and audit information, logging, and reporting. The maturity model evaluates where the organization stands by way of data governance readiness: *Level 1 Initial* operations are sporadic and rely on the knowledge of individuals' with respect to decision making; *Level 2 Managed* projects are administered but not from a cross-organizational perspective; *Level 3 Defined* standards are consistent across organizational units and individual projects; *Level 4 Quantitatively Managed* organizations set measurable quality goals leveraging quantitative techniques, and statistical metrics; and *Level 5 Optimizing* process improvement goals are established, and revised on an on-going basis to ensure process improvement.

Credibility. Berson holds a postgraduate degree in computer science. His professional interests and experience are in Master Data Management (MDM), customer data integration, and data warehousing working for Merrill Lynch, Dun & Bradstreet and PricewaterhouseCoopers, and others. He is a member of the Wall Street Technology Association's Board of Directors. Berson is published in technical magazines, and is the author and coauthor of professional books with a focus on data governance, data mining for CRM, data warehousing and client/server architecture. Dubov's professional focus is in Master Data Management (MDM), data warehousing, operational data stores, and Customer Relationship Management (CRM) with specific expertise in data and solutions architecture, and data stewardship. In addition to this book he is a coauthor of Master Data Management and Customer Data Integration for a Global Enterprise.

DAMA International. (2009). *The DAMA guide to the data management body of knowledge* (1st

ed.). Bradley Beach, NJ 07720. Abstract retrieved from

www.technicspub.com/books.htm.

Abstract.   The DAMA-DMBOK Guide is as a collective of standards and best practice for

data management professionals, executives, researchers and educators charged with the

responsibility to manage data and grow mature information infrastructures. More than 120 data

management practitioners contributed to this text.

Summary. Data is an essential enterprise asset. Managing data is a responsibility which must

be shared between IT data management professionals and business data stewards who represent

the interests of all data producers and consumers. Data stewardship designates business

responsibility, through a formal accountability process, to ensure optimal control and application

of data assets. As a business leader and/or subject matter expert, the data steward safeguards,

administers, and leverages data asset resources. "The best data stewards are found, not made…

responsibilities are not new and additional for these people. Whenever possible, appoint the

people already interested and involved" (p. 39). The appointment is a formal confirmation of

accountability in public recognition of their continued commitment. It is essential that data

stewards consider "the data interests of all stakeholders" and take an enterprise-wide perspective

on data assets, in order to guarantee "quality and effective use" (p. 39). Some organizations

differentiate between types of data stewards such as *executive data stewards* serving on the data

governance council, *coordinating data stewards* representing "teams of business data stewards",

and *business data stewards* recognized as subject matter experts (p. 40). There are three cross-

functional roles of stewardship and governance with judicial responsibilities: (a) *data

governance council members* have corporate-wide data oversight responsibilities consisting of

senior managers and executives representing both at the unit and enterprise level; (b) *data stewardship program steering committee members* provide support to the data governance council, preparing policies for consideration by the council; and (c) *data stewardship teams* business data stewards collaborating on data relevant activities within specific functional areas. This cohort relies on subject matter expertise to determine data definitions, quality requirements, and business rules. Data stewardship teams are standing, permanent units which meet regularly, frequently interacting with data architects.

Credibility. This book is written by a group of six members of the Data Management Association (DAMA) who formed the DAMA-DMBOK Guide Planning Committee. In addition, there were twenty-four international members on the Editorial Board, sixteen contributing authors, and eighty-two members who provided peer-review oversight.

Cheong, L. K., & Chang, V. (2007). *The need for data governance: A case study. 18th*

Australasian Conference on Information System. Paper presented at the 18th Australasian

Conference on Information System, Toowoomba, Australia. Abstract retrieved from

http://www.acis2007;usq.edu.au/assets/papers/171.pdf

Abstract. This paper investigates the concept of data governance as an emerging trend in

information management practices at the enterprise level. The relationship between data

governance and IT governance is explored. The measurements of data quality include accuracy,

relevance, completeness, timeliness, trustworthiness and contextual definition. Elements of

effective data governance and the adoption of various data governance models are addressed in

this work. The focus is on the collaborative side of the data governance framework specifically

between IT and business stakeholders.

Summary. Data governance is significant because it defines standards and procedures to

guarantee the proactive and effective handling of data management practices. The data

governance framework enables collaboration between multiple administrative roles to manage

data cross-organizationally and provides the capacity to align enterprise-wide objectives with

data related programs. In regard to the division of labor "the business' responsibility is to ensure

that the data is correct, available, reliable, and fit for purpose. IT is responsible for the

infrastructure that stores, processes and reports data… therefore, it seems logical that data

governance programs should be driven by the business side as business uses the data for

decision-making"(p. 1001). Cheong and Chang identify ten critical data governance success

factors as follows: (a) *strategic accountability* wherein the executive leadership drives data

governance processes and roles and responsibilities are clearly defined; (b) *standards* are

important since it is essential that enterprise data be refined and made fit for purpose; (c)

*managerial blind spot*s are avoidable when data is aligned with the proper technology,

operational processes, and governing bodies in conjunction with business objectives; (d)

embracing complexities of data while data producers and consumers collect, enrich, distribute,

and maintain information to share with various stakeholders; (e) *cross divisional issue* is

addressed when the data governance framework is designed to encourage participation from

every level in of the business in order to reconcile priorities and accelerate conflict resolution; (f)

data quality *metrics* is imperative in order to measure success of the data governance initiative;

(g) *partnership* with other companies requires the partnering organization to be equally

accountable for data quality to ensure that data management efforts are not undermined; (h)

strategic points-of-control determine the time and physical location of data quality assessments;

(i) *compliance monitoring* entails periodic assessment of data management policies and

operational processes; and (j) *training and awareness* of data stakeholders is invaluable to data

governance as an opportunity to promote the significance of high data quality.

Credibility. Cheong is a graduate of Curtin School of Information Systems where she

graduated with Distinction with a Masters in Information Technology. Over the course of her

career, Cheong has held a variety of roles in data analysis, applications management and

business analysis. Dr. Chang holds a PhD from Curtin Business School where she is the acting

head of the School of Information Systems. She is widely published in books, journal articles,

conference publications and presentations. One of Dr. Chang's main areas of research interest is

IT governance systems.

Gluck, K. (2008). Information stewardship: Roles and responsibilities in Vattenfall's

maintenance process. *Masters' Thesis, ETH Electrical Engineering.* Retrieved from

http://www.isaca.se/dynamaster/file_archive/091216/3d6026d5e367614f125b0bb3d127dec8/

Uppsats.pdf

Abstract. In recognition of data as "one of the most valued assets… the structuring of

information" is a crucial matter for organizations (p. ii). This masters' thesis examines

Vattenfall's governance structure within the Enterprise Informational Architecture. Roles and

responsibilities specific to the information maintenance process provide the primary focus of this

study.

Summary. The premise behind the evolution of stewardship is that information is not the

property of an individual but rather that of the organization. This concept has led to the

assignment of stewardship roles and responsibilities which consider information from a strategic,

operational or tactical level. Stewardship provides the accountability essential for clearly

defining, creating, modifying, distributing, deleting, optimizing and compiling information.

People are responsible for information in every organization; a problem exists when "the

accountability is not formalized" (p. 22). In addition, people often view themselves as the owners

of information passing through their systems, which creates the opportunity for multiple

definitions of the same data across divisional and multi-platform boundaries. Stewardship, as

defined by Merriam-Webster's online dictionary is "the careful and responsible management of

something entrusted to one's care…" which is different from ownership.

When assigning roles and responsibilities, data handling can be broken down into three

actions: (a) those that define information, (b) those that produce information, and (c) those that

use information. One key requirement is that the role of stewardship should not be restricted to

IT staff. It is important that the data is managed by the business side of the house in cooperation with IT. "If only IT drives the [data governance] program the organization is doomed to fail" (p. 23). This is primarily because the business understands the data in organizational context. Another important requirement is that every business function throughout the organization should be assigned an information steward thus ensuring accountability at the enterprise level. Stewards should be influential, visible and respected with a good understanding of the organization's vision and the ability to communicate it to others. Additionally the stewards should have senior level support. The most important requirement is a strong organizational culture backing the information stewardship program with formal accountability measures.

Credibility. Gluck is a graduate student with the KTH Royal Institute of Technology in the School of Electrical Engineering. Gluck acknowledges both Pontus Johnson and Bjorn Ekstedt as instrumental in support of this master's degree research. Johnson, one of Gluck's professors's at the KTH Royal Institute of Technology, heads the Department of Industrial Information and Control Systems, and has published many research papers on enterprise architecture analysis. Ekstedt works at Vattenfall, the organization providing case study research for this work. He heads the department of Operations IT Supply and is also self-employed as a consultant of Strategies in Leadership Development.

Khatri, V., & Brown, C.V. (2010). Designing data governance. *Communications of the ACM*,

53(1), 148-152. Retrieved from

http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=true&

db=buh&AN=47220627&login.asp?custid=wchstcc&site=ehost-live

Abstract. The article makes a distinction between data management, and data governance.

The need to comply with regulatory pressures has led many organizations to re-examine and

revalue their data handling processes from a strategic perspective. Data governance acts as the

foundation for developing a decision matrix to determine the appropriate level of responsibility

surrounding data handling practices. The work presents a guideline with five key zones of

decision-making for implementation of a data governance stratagem.

Summary. This article provides a data governance framework for information management

practitioners. This approach, originally designed by Weill and Ross for IT governance, is

modified in this article to present a relevant alternative strategy for data governance. The design

includes five key decision domains: 1) *data principles* establish how all other decisions about

data assets will be conducted. Data principles set the boundaries for intended use, determining

the organization's stance on data quality, and leading to how data is defined (metadata) and how

it is accessed, and by whom. Key decisions in this domain may define desirable behaviors for

using data as assets, and opportunities for sharing, as well as addressing the regulatory

environment impacting data usage; 2) *data quality* establishes standards in regard to accuracy,

credibility, completeness, and timeliness. Key decisions in this domain may inquire on the

protocol to determine data quality, how to communicate on the topic, and establishing program

evaluation procedures; 3) *metadata* is simply defined as "data about data" (p. 150) establishing

the rules to interpret data. There are different types of metadata. *Physical metadata* defines

where data is stored, domain-independent metadata provides a descriptive audit trail of

information about the data such as the creator or modifier of the data and authorization and

lineage of the data. *User metadata* provides annotations on things such as user preferences, and

the history of usage. Key decisions around metadata might include the plan for updating different

types of metadata, and the approach for consistently defining data to ensure it is deciphered

correctly; 4) *data access* refers to the standards set by an organization's access policies and may

integrate audit tracking, privacy and availability practices. Data integrity standards, for example,

ensure data is safe from physical damage that could occur as a result of an unanticipated power

failure. Conversely, logical data integrity is designed to protect the database structure. Key

decisions about data access might include determining the business value of data, establishing

ongoing risk assessment measures, designing standards and procedures for data access,

monitoring periodically for compliance, deciding upon methodologies for communicating and

educating on matters of security, and developing backup and recovery programs; 5) *data*

lifecycle refers to all of the stages from creation, usage, storage to deletion that a data element

undergoes. "By understanding how data is used, and how long it must be retained, organizations

can develop approaches to map usage patterns to the optimal storage media, thereby minimizing

the total cost of storing data over its life cycle" (p. 151). In addition, within each the five

decision domains are the assignment of potential roles for accountability and decision-making

purposes.

   Credibility. This article is published in a peer-reviewed journal. Khatri is an Associate

Professor of Information Systems at Kelley School of Business at Indiana University. He holds a

PhD with the University of Arizona. He is frequently published, writing articles associated with

information systems, data modeling and knowledge management. Brown is a Distinguished

Professor and Director of Healthcare IT Management at Howe School of Technology

Management. She holds a PhD in Management Information Systems. Brown is published

extensively in professional journals, conference proceedings, reports, and books; most notably

MIS Quarterly Executive, Journal of Management Information Systems, Information Systems

Research, and Information Systems Management. Both Khatri and Brown are recipients of

professional awards and honors for best paper, and teaching excellence.

Redman, T.C. (2005). A comprehensive approach to data quality governance. *A Navesink*

Consulting Group White Paper. Retrieved from http://iaidq.org/webinars/doc/2006-11-

redman-data-governance-white-paper-2005-06-22-version.pdf

Abstract. This paper presents a comprehensive strategy from which to approach data quality

governance. While most companies wholeheartedly agree on the idea of data as key assets, they

are just as likely to admit they are in bad shape. "Too often the data needed for critical decisions

and operations are unavailable, poorly defined, out-of-date, incorrect, or otherwise unfit for use.

Further, traditional hierarchical organizational structures are ill-suited to managing data" (p. 1).

Summary. The most crucial issue that data governance must tackle is to ensure the proper

management accountability. This model does so through the six jointly-reinforcing components:

1) The *data council* relies on senior leadership. The data council represents the leadership of

the organization. "The seniority and position of the individuals perceived to be leading the data

quality effort dictate its success more than any factor" (p.8). As a group they: (a) determine the

purpose of the entire data quality effort; (b) decide upon data-related policies which include the

assignment of roles and responsibilities, identify data ownership, data sharing, and privacy; (c)

are responsible for the data supplier structure and the information chain; and (d) reinforce the

data quality culture and fund data quality training.

2) The *data quality staff* is led by a Chief Data Officer who reports to the Data Council's

most senior member. The Chief Data Officer, leading a small full-time staff of employees, is

responsible for designing the data quality strategy, and measures the success of the effort. This

group manages meta-data and operational processes.

3) *Information chain management* refers to the governing body that provides oversight of

internal data which includes input from suppliers, follow through on the process to manipulate

data, outputs to customers, and the collection and analysis of feedback along the way. The information chain management is responsible for assembling "a team of managers, drawn from various functional areas... one especially senior manager, the "owner", is assigned to lead the team. Assembling this team, naming the "owners" and ensuring responsibilities and decision rights are clearly defined as the responsibility of the information chain management cycle" (p. 13).

4) *Data supplier management* is similar to the information chain management but deals with external data sources. This group assigns supplier management areas of responsibility and engages selected suppliers. They develop customer requirements, provide a baseline for supplier performance, and plan, control and improve the data handling process through performance tracking processes.

5) S*pecific information technology roles* define IT's area of data quality responsibility which extends to the technical infrastructure and includes the organization's databases, communication infrastructure, and all computer applications. This team is accountable for the implementation of security and data policies as they relate to privacy. It also presents technology applications that align with business needs, ensures access to the right people, provides tools to build high-quality data applications, and implements data clean-ups.

6) Chartered *improvement team* provides continuous improvement to processes on a project-by-project basis.

Credibility. Redman is a trained statistician starting his career over thirty years ago working for Bell Labs where he focused his attention on data quality, a good ten years before anyone acknowledged there was an issue. In 1996 he established his own business, Navesink Consulting Group, assisting organizations in a data management capacity as an experienced data quality

consultant. Redman is quoted extensively in data quality circles. He has a PhD in statistics from

Florida State University.

Soares, S. (2011). The IBM data governance unified process: Driving business value with IBM

software and best practices. *MC Press Online*, LLC, Ketchum, ID.

Abstract. This step-by-step guide is based on IBM data governance best practices. This

manual includes tools and strategies to treat data as corporate assets. Topics include the

optimization of decision rights, and securing and leveraging data. This text provides a fourteen

step data governance framework addressing core issues focused on people and processes.

Summary. Data governance bestows "decision rights to optimize, secure, and leverage data

as an enterprise asset" (p. 3). The IBM data governance unified process incorporates a fourteen

step program: designed to: (a) *define the business problem*; (b) *obtain executive sponsorship*

from major IT and business leaders; (c) *conduct an annual maturity assessment*; (d) *build a

roadmap* to link the current state of corporate maturity with the preferred future state of maturity

for each of the maturity assessment categories; (e) establish an *organizational blueprint* to

govern operations, and ensure authority; (f) *build a data dictionary,* a type of glossary, from

which to define key business terms; (g) make sure the *data is understood*; (h) create a *metadata

repository* to define data characteristics; (i) *define metrics* to measure performance and track

progress; (j) *govern the master data* of information which is business-critical; (k) *govern

analytics* enabling an alignment between business users and investments analytics; and (l)

manage security and privacy (see figure 3).

Stakeholders should consider the eleven elements below when conducting an annual maturity

assessment:

1. *Data risk management compliance* which is a method where "risks are identified,

    qualified, quantified, avoided, accepted, mitigated, or transferred out" (p. 31);

2. *Value creation* deals with qualifying and quantifying data assets to enable the enterprise to exploit the value produced by data assets;

3. *Data risk management compliance* which is a method where "risks are identified, qualified, quantified, avoided, accepted, mitigated, or transferred out" (p. 31);

4. *Value creation* deals with qualifying and quantifying data assets to enable the enterprise to exploit the value produced by data assets;

5. *Organizational structures and awareness* has to do with the level of shared responsibility among IT and business, and the acknowledgment of fiduciary duty to govern data assets at numerous levels of management;

6. *Stewardship* refers to a quality-control discipline which ensures "custodial care of data for asset enhancement, risk mitigation, and organizational control" (p. 31);

7. *Policy* is written to articulate the desired organizational behavior;

8. *Data quality management* is the measurement process used to improve, and confirm the quality, and reliability of test, archival, and production data;

9. *Information Lifecycle Management* is the policy-based method used to systemically compile, use, store and delete data assets;

10. *Information Security and Privacy* has to do the organization's controls, such as policies and standard practices that are used to reduce risk and secure data assets;

11. *Data Architecture* "is the architectural design of structured and unstructured data systems and applications that enables data availability and distribution to appropriate users" (p. 32);

12. *Classification and metadata* provide tools and methods to create common definitions for data models, IT and business terms, and repositories;

13. *Audit Information Logging and Reporting* is the organizational process to monitor
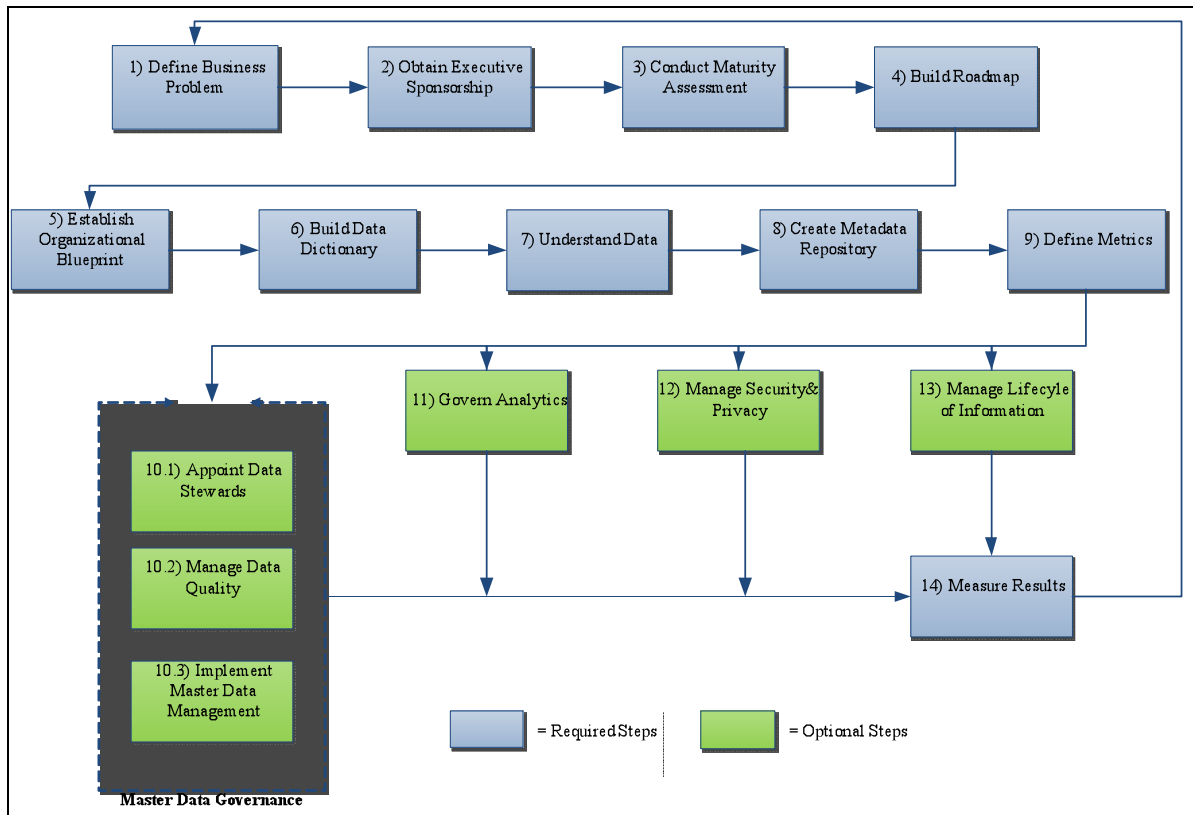
and gauge the programs value, and risks;

Figure 3. IBM Data Governance Unified Process Overview (p. 8).

Credibility. Soares has an MBA from the University of Chicago. He is employed with the

IBM Software Group as Director of Data Governance as part of a team of more than 200

professional consultants who assist IBM clients in the assessment of organizational maturity

levels on data governance practices, and work with clients on the integration of appropriate data

governance processes and tools.

Thomas, G. (2009). How to use the DGI data governance framework to configure your program.

The Data Governance Institute. Retrieved from

http://www.datagovernance.com/wp_how_to_use_the_dgi_data_governance_framework.pdf

Abstract. The data governance framework is a logical way of organizing, classifying, and communicating intricate decision making activities surrounding the use of corporate data. The purpose of this work is to assist business, IT and data management professionals with rules, people, tools and processes having to do with the governance of enterprise data.

Summary. The DGI framework is based on the premise that businesses have direct information needs, which guide technology strategies. "To succeed in this mission, technology teams must understand those information needs… they need definitive input from business resources" (p. 4). The data governance framework (see figure 4) provides a tool to determine and enforce rules of the program (such as policies, standards, controls, definitions, etc.), and rules of engagement which describe how various groups collaborate to create and enforce rules. A second schema within the framework addresses some simple data-related questions:

Why should the program exist?

What will the program accomplish?

Who will be held responsible through the designation of specific accountabilities?

How will group collaborate to achieve value on behalf of the organization?

When will particular processes be performed?

Both schemas are shown in the DGI data governance framework in figure 4 below:
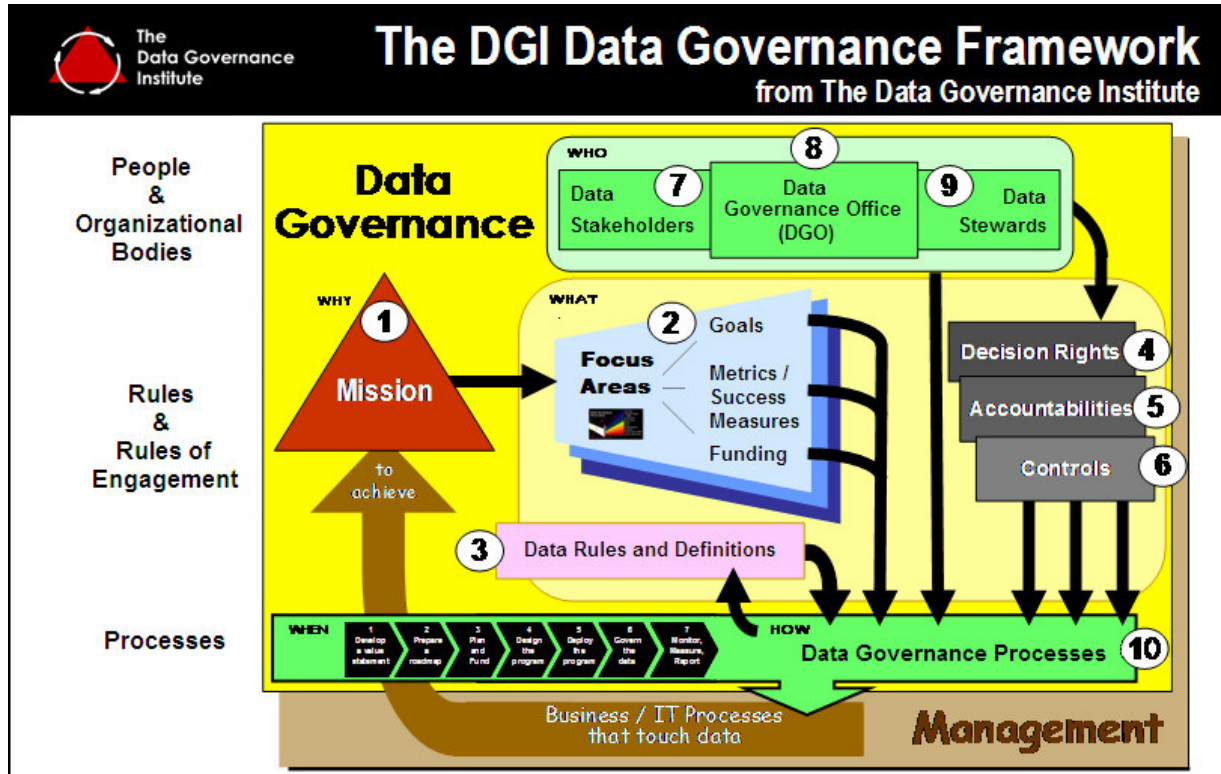
Figure 4. The DGI Data Governance Framework (p. 6).

The framework also relies on a program lifecycle (see figure 5). Following the lifecycle of the program affords clarity around the particular business problem being addressed, and drives monitoring activities to achieve accountabilities.
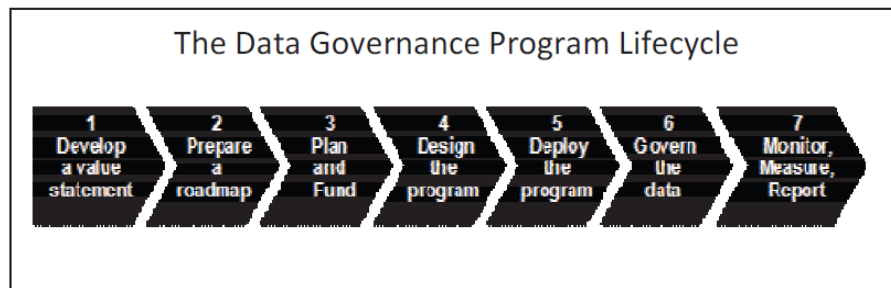

Figure 5. The Data Governance Program Lifecycle (p. 7)

Credibility. The Data Governance Institute provides vendor-neutral information on data governance and stewardship tools and best practices; supplying techniques and models, consulting services, and training. The institute also houses the site for The Data Governance Community of Practice containing a repository of free program documents and case studies for information exchange supplied by other practioners/members. Thomas is the founder of the Data Governance Institute where, as acting president, she hosts the website and manages its membership. This article is published by the Institute. Her career spans twenty years specializing in data management and governance, systems integration and other related areas. Thomas is a key-note speaker at international conferences and symposiums.

Villar, M. (2009). Establishing effective business data stewards. *Business Intelligence Journal*,

14 (2) 23-29. Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=415

22172&site=ehost-live

Abstract. Still an emerging role, business data stewards are responsible for data definition

and consistency, accuracy, and timeliness of critical information. Driven by financial reporting

requirements, federal privacy mandates, and the need for security increases are part of the force

behind the implementation of data governance programs causing a growing demand for

companies to establish business data stewards. Most organizations are struggling with how to

effectively implement a stewardship program as the role is still emerging. This article addresses

the role and responsibilities of the business data steward through the outline of successful

attributes.

Summary. The business data steward is responsible for oversight of the critical business data

including newly created data, and data being used, stored within the particular stewards'

functional scope. "Not all data requires a business data steward, but certainly all critical data

should be assigned one" (p, 24). The enterprise data executive, chief data officer or chief

operating officer is responsible for oversight of the stewardship program. Business stewardship

should be "high on the staff hierarchy and as visible as possible to drive the data programs across

multiple organizations" (p. 24). The role of business steward should be tailored to the

organization's level of information management maturity, institutional culture and data issues.

Business stewardship is a leadership role; someone who is directly impacted by data quality in

his or her position. Business stewards interpret strategy transforming them into tactics to achieve

business goals. The steward identifies high-value critical data that meet the needs of the business.

They ensure that data meets certain standards, driving quality improvement and effective

business processes. Data accessibility and archive management fall within their purview.

"Stewards drive the consolidation and simplification of the database landscape. When business

managers want to build or acquire a new database, they should first seek the data steward's

advice. The task of consolidating redundant, underused databases falls on the business data

steward working in concert with IT teams" (p. 27). Business data stewards stimulate change in

processes, technology, and governance to guarantee that business data goals can be met. While

the executive data steward is responsible for the entire stewardship program, the business data

steward is accountable for a particular domain area. Although the executive data steward is not

automatically a full-time role the business data steward should be.

Credibility. Villar is a 25 year veteran in IT, enterprise data management and technology re-

engineering. She has held executive positions at the senior level with responsibilities in both data

quality and governance. Her work creating the first corporate-wide "Enterprise Business

Information Center of Excellence at IBM" was recognized by an external entity, The Data

Warehousing Institute (TDWI), as a data governance best practices business intelligence

application. In addition, she has received national awards for her work.

Weber, K., Cheong, L.K., Otto, B., & Chang, V. (2008). Organizing accountabilities for data

quality management – A data governance case study. *Institute of Information Management,*

University of St. Gallen, Muller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland. Retrieved

from

http://web.iwi.unisg.ch/org/iwi/iwi_pub.nsf/wwwPublYearGer/028D5245B5975BF7C12574

F1002EBB0E/$file/DW2008%20FP%20Data%20Governance%2001%20kwe.pdf

Abstract. Businesses need to manage data quality at the enterprise level combining

business-driven perspectives with technical aspects in order to respond strategically to

operational challenges. However, companies tend to assign accountability for data quality

management to their IT departments. In doing so, they fail to recognize organizational issues

inherent to data quality management success. Data governance can address this issue with

corporate-wide accountabilities thereby bridging the gap between business and IT.

Summary. This article focuses on the accountabilities side of data governance. The

autocratic top-down structure of organizations can present barriers to successful data quality

management when data is used across the enterprise; data governance applications can help to

traverse these obstacles (Thomas, 2006) by supporting global operations and regulatory

compliance. Organizations have a propensity to assign IT departments with the responsibility for

data quality improvement and management practices wherein in order to tackle both

organizational issues and the IT point of view, an integrated data quality management approach

is called for. In *governance by management* decision-rights are executed by managers. Although

managers are responsible for establishing data-related rules there is no formal assignment of

responsibilities. *Governance by stewardship* takes a more formal approach by assigning roles

and responsibilities. *Governance via governance*, the more common approach, distinguishes

between those that govern, making rules and resolving issues, and data stewards who work with the data, ensure rules compliance and manage issues.

Governance bodies complement corporate-hierarchical structures enabling them to deal more effectively with cross-organizational data management concerns. The *governance by governance* approach defines five primary roles as the organizational governing structure: 1) an executive sponsor adds a strategic dimension, advocacy and funding to the data governance program; 2) the chief data steward is the enforcer of standards, and assists in determining metrics and objectives; 3) the business data steward, working within his specific area of responsibility, implements data quality standards and policies; 4) technical data stewards work at the data element level dealing with definitions and formats, examining aspects of both source systems and data flow operations between systems; and 5) the data quality board is chaired by the chief steward and determines enterprise-wide and controls.

Credibility. Weber is a Research Assistant at the University of St.Gallen's Institute of Information Management in Switzerland where she holds a doctorate of information management through the same institution. Her primary professional focus is on data quality management through corporate data governance. Cheong is a graduate of Curtin School of Information Systems where she graduated with Distinction with a Masters in Information Technology. Over the course of her career, Cheong has held a variety of roles in data analysis, applications management and business analysis. Dr. Otto is an assistant Professor, head of Corporate Data Quality with the University of St. Gallen. He is published in scientific journals, conference proceedings, books, and invited presentations. Dr. Chang holds a PhD from Curtin Business School where she is the acting head of the School of Information Systems. She is

published in books, journal articles, conference publications and presentations. One of Dr.

Chang's main areas of research interest is IT governance systems.

Welch, M. (2009, June). *Information governance and stewardship for records and information management*. Abstract retrieved from

https://www.cis.unisa.edu.au/wki/images/5/58/Welch_thesis_June_01_2009.doc.

Abstract. This masters' thesis examines the theory that information would be better managed if business stewards were held accountable for ensuring data quality. Two primary conclusions are presented: 1) information management professionals could leverage governance and stewardship practices for this purpose, and 2) data could be managed holistically from a cross-organizational perspective.

Summary. A primary goal of the data governance program is to ensure high quality information assets throughout the lifecycle of data. "Data governance includes a formal process known as stewardship that is key to ensuring information quality. Data stewards manage organizational data on behalf of the organization and its staff" (p. 11). The generic stewardship model (see figure 6) depicts the levels of responsibility of stewards as sponsors of information assets. It facilitates widespread collaboration on multiple levels throughout the organization. This model fosters an alliance between IT and business units providing oversight and advise on strategy, both on tactical and operational levels. It incorporates rules and procedures, as well as compliance and enforcement opportunities. Executive buy-in at the sponsorship level is critical to success of the program. The information governance council operates at a high level of executive sponsorship. The council directs and monitors stewardship groups and the groups monitor individual stewards. It is important that the council consist of senior level employees with firm big-picture understanding of the business in addition to representing specific areas of the organization. Domain stewards are on the level directly below the information governance council. This group bridges the gap between IT and business and is held for data quality within

specific information domains. "It is critical that stewards at this level represent all areas of

business that have an interest on the quality and integrity of the information with that domain"

(p. 56). Theoretically, the lowest level of stewardship represents everyone else in the

organization. "All staff must accept some level of responsibility for the accuracy and quality of

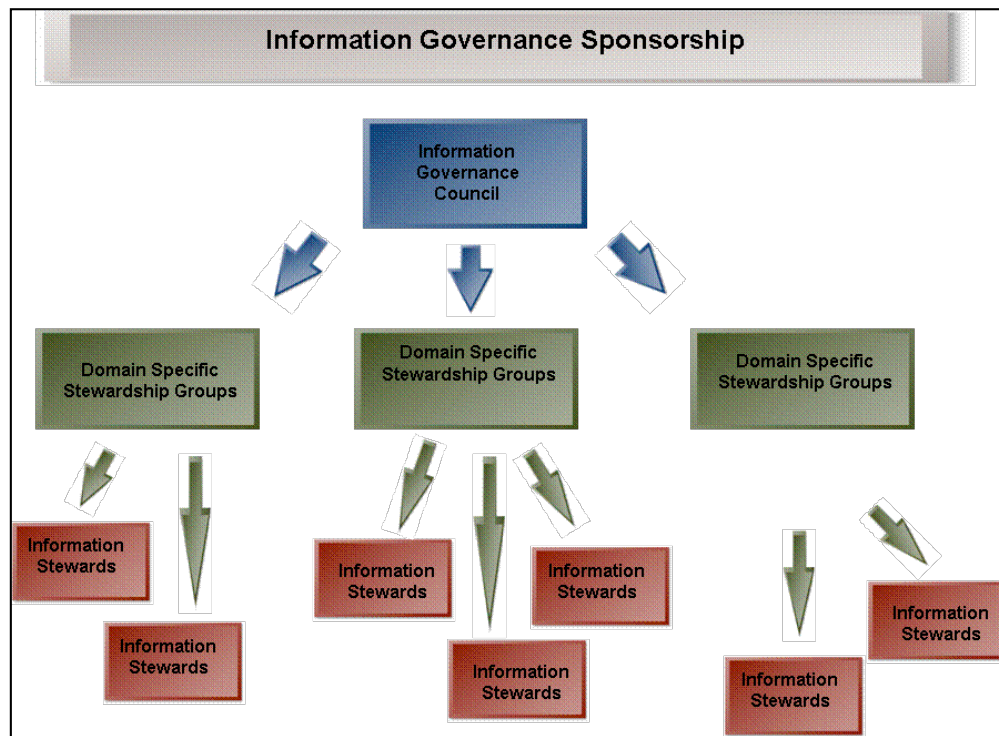information and data they deal with on a day-to-day basis" (p. 57).

Figure 6. Generic Stewardship Model (p. 54).

Credibility. Welch presents this master's thesis through the University of South Australia

(UniSA). She is the Director of Profile Records Management Services Limited and has been

working as a consultant in records management since 1986 in both New Zealand and Australia.

In 1990 she developed a course in records management at Auckland's Institute of Technology; a

class which she also teaches. Welch was the president and founder of the Chapter in Auckland

for the Association of Records Managers and Administrators (ARMA) International.

Wende, K., Otto, B. (2007). A contingency approach to data governance. *University of St. Gallen, Switzerland*. Retrieved from

http://mitiq.mit.edu/iciq/Documents/IQ%20Conference%202007/Papers/A%20CONTINGE NCY%20APPROACH%20TO%20DATA%20GOVERNANCE.pdf

Abstract. Historically, organizations have assigned data quality management accountabilities primarily to IT departments, thereby ignoring business issues crucial in the successful management of data quality. Conversely, data governance assigns corporate-wide accountabilities spanning business and IT. This paper presents a data governance framework containing three components built on a comparable matrix to a RACI (responsible, accountable, consulted and informed) chart.

Summary. Data governance is seldom adopted according to a 2005 survey of 750 participants, with only 8% indicating they had deployed a program (Russom, 2006). In the meantime, companies seek competitive advantage by focusing on IT investments when successful data quality initiatives require collaborative efforts between IT and business specialists. This model of data governance (table 1) consists of "data quality roles, decision areas and main activities, and responsibilities, i.e. the assignment of roles to decision areas and main activities" (p. 6). "The data governance model uses a set of four roles and one committee-the data quality board" (p. 7). Roles include the Executive Sponsor, Chief Steward, Business Data Steward, and Technical Data Steward in addition to the Data Quality Board. Table 1 provides a brief description of decision areas for each role. A data governance framework helps organizations in structuring accountabilities around data quality initiatives. The matrix depicts roles of data governance and the degree of authority within each role. Each data governance

configuration is unique for each organization. Once data quality roles are defined, decision areas

are identified, and responsibilities delegated, components are arranged in the data governance

model as an outline of the data governance structure.

Roles Decision Areas	Executive Sponsor	Data Quality Board	Chief Steward	Business Data Steward	Technical Data Steward
Plan data quality initiatives	A	R	C	I	I
Establish a data quality review process	I	A	R	C	C
Define data producing processes		A	R	C	C
Define data roles and responsibilities	A	R	C	I	I
Establish policies, procedures and standards for data quality	A	R	R	C	C
Create a business data dictionary		A	C	C	R
Define information		I	A	C	R
R = Responsible; A = Accountable; C = Consulted; I = Informed					

Table 1. Illustrates the draft of a data governance model (p. 6).

Credibility. Weber is a Research Assistant at the University of St.Gallen's Institute of

Information Management in Switzerland where she holds a doctorate of information

management through the same institution. Her primary professional focus is on data quality

management through corporate data governance. Otto is an assistant Professor, head of

Corporate Data Quality with the University of St. Gallen. He is published in scientific journals,

conference proceedings, books, and he has delivered invited presentations.

Wende, K. (2007). *A model for data governance – Organizing accountabilities for data quality management.* Paper presented at the 18th Australasian Conference on Information System, Toowoomba, Australia. Retrieved from http://aisel.aisnet.org/acis2007/80.

Abstract. Data quality management (DQM) at the enterprise level integrates business and technology requirements to respond strategically to business challenges demanding high quality enterprise data. Up until now, responsibilities for DQM fell primarily to IT, ignoring the inherent organizational issues essential to DQM success. Conversely, data governance provides enterprise-wide accountabilities bridging the gap from business to IT. This paper provides the framework for a data governance model containing three components.

Summary. There is a prevailing assumption that a universal approach to data governance is effective as a one-size-fits-all data management solution. However, Wende suggests that such a practice isn't effective since it lacks "an elaborate analysis of the interaction of roles and responsibilities, and the design of decision-making structure" (p. 418). Hence, organizations might find it challenging to sustain data of high-quality at the enterprise level. Achieving corporate data quality necessitates collaboration between various business and IT stakeholders with a firm grasp of the data and its purpose.
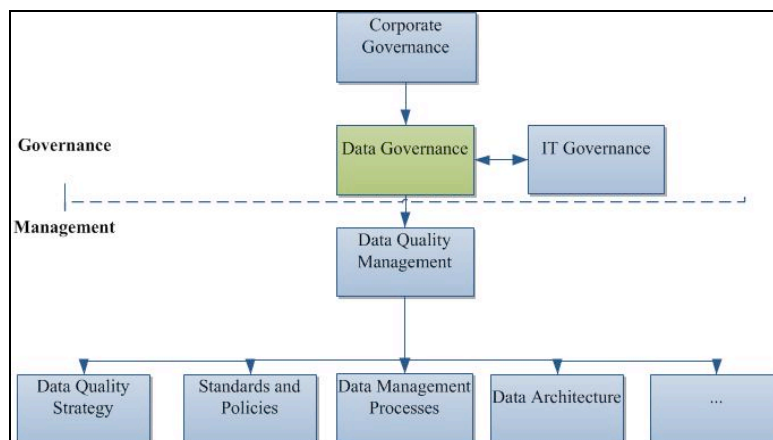

Figure 7. The Data Governance Matrix (p. 419).

Figure 7 shows the relationship between data governance and other areas of data quality management (DQM). The data governance framework incorporates DQM roles, areas of decision making, core activities, and responsibilities. Specific configurations are unique to each company but more common components include establishing the definitions of data quality roles, areas of decision-making, and assignment of responsibilities. Although the actual quantity may vary, the following four sample roles and single committee are provided: 1) *the executive sponsor* provides strategic direction, oversight, funding, and advocacy; 2) the *chief steward* enforces the standards established by the data quality board and identifies metrics and targets related to data quality. In addition, another important role of the chief steward is to chair the data quality committee; 3) the *business steward* documents business requirements and appraises the impact of data quality requirements. They also communicate with the data quality board on the recommendation of standards and policies from the business perspective; 4) the *technical data steward* equivalent of the business data steward and are responsible for data representation from the IT perspective; and 5) the *data quality board* outlines the data governance structure at the corporate level, setting strategic goals, developing and directing enterprise-wide policies, rules, standards, operational guidelines with the objective of improving data quality.

Credibility. This author appears in numerous publications on the topic of data governance. Wende is a Research Assistant at the University of St.Gallen's Institute of Information Management in Switzerland where she holds a doctorate of information management through the same institution. Her primary professional focus is on data quality management through corporate data governance.

Zornes, A. (2006). Corporate data governance best practices. *The CDI Institute MarketPlus TM*

In-Depth Report. Retrieved from

http://www.tcdii.com/PDF/Data_Governance_white_paper_-_April_2006.pdf

Abstract. This white paper is based on the survey responses of 54 IT executives from Global

5000 companies; the largest 5000 institutions in the world. Participants include chief IT

management and business data stewards. The following industries are represented in this survey:

financial services providers, high technology manufacturers, communications services providers,

and representatives from the hospitality industry.

Summary. Zornes describes data governance as "the formal orchestration of people,

processes, and technology to enable an organization to leverage data as an enterprise asset" (p.

1). The challenge of dealing with data today far outweighs the information management demands

of the past due to the sheer volume and complexity of recent enterprise assets. Companies today

need to integrate data spread across entire divisions. Data governance requires businesses to

dissect information silos, and assimilate staff cross-organizationally from within different

functional areas, lines of business, and geographic regions. As an evolving entity data

governance integrates centralized [data handling] policies. "…best practice within companies

successfully implementing data governance is the collaboration between IT management and

business leadership to design and refine *future state* business processes associated with data

governance commitments" (p. 2). Furthermore, a robust data governance application is integral

to the preservation of reliable, usable business data. An additional best practices approach is to

appoint data stewards or a mixture of business unit stewards and corporate data stewards to

exercise quality oversight of data assets. Business benefits include: (a) operational savings and

increased efficiencies; (b) privacy and compliance; (c) consistent customer treatment; (d)

infrastructure from mergers and acquisitions; and (e) enhancement of revenue and customer loyalty.

Credibility. Zornes is founder of the Customer Data Integration (CDI) Institute where he operates as the chief research officer. He specializes in master data management (MDM) and customer data integration (CDI) and is the editor and chief contributor of the CDI Newsletter published in the DM Review. This white paper is published by the CDI Institute. Zornes holds a MS in Management Information Systems through the University of Arizona.

Examples of Data Governance Practices in Selected Institutions of Public Higher Education

Albrecht, B. & Pirani, J.A. (2009). Revitalizing data stewardship through risk reduction:

Managing sensitive data at the University of Virginia**,** *EDUCAUSE Center for Applied*

Research, Case study 8, retrieved from http://www.educause.edu/ecar

Abstract. This case study explores the ways in which higher education institutions deal with

institutional data management challenges in terms of content, records management, quality,

stewardship, governance, research data management and analytics. Data was compiled from 309

web survey responses distributed primarily to senior IT leaders from 1,733 EDUCAUSE

member organizations, and follow-up telephone interviews were conducted with 23 institutions

for more in-depth analysis. The study was performed to expand and improve upon the findings of

an earlier study on *Institutional Data Management in Higher Education* which dealt with

practical solutions to sensitive data issues.

Summary. Many colleges and universities are witness to the dramatic growth rate of

institutional data influenced in some degree by the influx of technology. Personal computers,

laptops, note pad devices, smart-phones, and the like foster quick and easy access to data

anytime and anywhere. Storage media is the size of a thumbnail, can be carried around on a

keychain, and collected in seconds. "All these factors make the care of institutional data an

increasingly vital and complicated process, requiring policies and processes to ensure security,

accuracy, and timeliness, as well as accessibility and readability—all of which a conceptualized

sometimes under the name of stewardship" (p. 2). Several drivers lead to the development and

implementation of a data governance program at the University of Virginia. The first was anxiety

expressed by Vice President Hilton in regard to risk management and mitigation, and

surrounding the institution's handling of sensitive data. Second, the college community began to experience unease around data handling practices. The third driver was the rapid proliferation of decentralized data environments. "The regulatory environment also prompted a reassessment of data handling practices" (p. 4). The university has both an administrative data access policy and an information technology security risk management program; the latter incorporates a multi-layered strategy for the implementation of data stewardship, which is driven by privacy, security, data retention requirements, and data access procedures.

Credibility. Albrecht has held the following academic roles: faculty member, senior fellow, dean, vice provost, academic affairs vice president, deputy commissioner for academic affairs, and chancellor emeritus within multiple institutions: including the University of Chicago, the University of Oregon, the University of Northern Colorado, and the Montana University System. He has published on the topics of technology-supported learning and distance learning. Pirani is a Fellow at the EDUCAUSE Center for Applied Research (ECAR). She conducts higher education research and analysis focusing on IT-related issues. Pirani is the author of more than 30 ECAR supported case studies examining leading-edge technologies and identifying exemplar management practices.

Drucker, P. (2005). Academic analytics in the future of higher education. EDUCAUSE Center

for Applied Research. Retrieved from

http://net.educause.edu/ir/library/pdf/ers0508/rs/ers05089.pdf

Abstract. This EDUCAUSE Center for Applied Research (ECAR) Study focuses on the

challenges of academic analytics while higher education institutions in the U.S. face a

tumultuous economy and rapidly changing political environment.

Summary. Many liken the political and economic conditions facing U.S. higher education to

the *perfect storm*. Changing financial markets, growth in noncredit instruction presenting new

competition, andd a handful of highly-sought-after private institutions leave many universities

struggling to increase retention, battling over limited student dollars. Many believe that the

educational landscape is changing... demanding greater accountability and transparency from

higher education" (p. 3). As a result, college and universities will be seeking new opportunities

to stimulate growth: focusing on new revenue sources, placing more emphasis on "time-to-

market issues and hence practices that affect the velocity of decision making" (p. 4). Executive-

level administrators are relying on data with more frequency and "may become impatient when

access to comprehensible information or sophisticated analysis is limited, constrained, or

nonexistent" (p. 6). Data governance, the tool used to manage essential information, is expected

to be a challenge in this environment. Information is often dispersed among units, departments or

divisions and thereby subject to varying standards and access policies for example. Integrating

data housed in silo systems will require a federated approach to data management and new

technologies.

Credibility. Peter Drucker is a teacher, writer, and business strategy and policy consultant.

Drucker is the recipient of multiple world-wide honorary doctorates. He has authored thirty-one

books and numerous articles. His works appear in the Harvard Business Review, the Wall Street

Journal, and EDUCAUSE which promotes the intelligent use of information in higher education.

Frost, J., Lucas, C., & Blankert, J. (2004). Strategies and tools used to collect and report

strategic plan data, Online Submission. Paper presented at the *Annual Forum of the*

Association for Institutional Research (AIR) (44th, Boston, MA). Retrieved from

http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED491018

Abstract. This paper covers the process to foster cross-organizational stakeholder

relationships, including the collection and oversight of data along with the configuration of facts

for straight-forward presentation to diverse populations. With the swell of both internal and

external data accessible to higher education institutions, and the proliferation of technical tools

available with which to manage and access data, institutions have the chance cultivate

knowledge management. This work presents strategies and tools to aide in strategic planning

data compilation and reporting.

Summary. In 1999 Purdue University implemented two critical knowledge management

activities: 1) an Institutional Data Network (IDN) was formed as a group charged with: (a)

ensuring the accuracy and consistency of data sent to external sources; (b) developing a

collective voice in response to information requests; and (c) functioning as a judicial body to

supply accurate institutional research data. A wide cross section of 45 staff members

representing 20 different offices sat on the IDN meeting monthly to discuss data-related issues;

and 2) a data digest was designed comprising of historical university information on students,

faculty and staff, instruction, facilities and research. "In addition to now having a single

authoritative source of data about the university, the climate for understanding the need for

standard definitions, authoritative sources, and data experts had been established" (p. 3). Purdue

began to scrutinize their data handling practices in August of 2000 prompted by a new president

with a data-driven decision-making management style. In light of new technologies and the ease

of access to global data sources, universities face an escalating crisis of too much information, combined with an abundance of unreliable information, and improperly stored, or not easily accessed data (as cited in Teodorescu & Frost, 2002). Historically, IR Offices have functioned as repositories for information, gathering, storing, and formatting data for other institutional offices. "However, these positions, as effective and necessary they are, create a disjuncture between the collection of data and the use of data in decision-making. If data in our organizations function to increase intelligence, inform policies and aid planning, the information must be tied to the audience who will use it and the need it will fulfill" (p. 4).

Credibility. Frost is the Director of Institutional Research with Purdue University where he is responsible for governance reporting, national survey results and peer benchmarking. Lucas is a Research Analyst at Purdue University. In 2008 she was a presenter of a workshop on IPEDS comparison tools at a regional conference in Dearborn, Michigan. Blankert is a graduate assistant, also with Purdue University. The article is published in conference proceedings.

Yanosky, R. (2009). Key findings: Institutional data management in higher education.

Educause. Retrieved from http://net.educause.edu/ir/library/pdf/EKF/EKF0908.pdf

Abstract. This Educause Center for Applied Research (ECAR) study examines policies and

practices from which higher education institutions successfully compile, secure, and employ

electronic information assets to address business needs, and satisfy academic requirements. This

study includes a literature review investigating various data management structures and

definitions; Yanosky worked in consultation with IT administrators in higher education as well

as experts in the field of data management. More than 300 EDUCAUSE member institutions

responded to a web-based quantitative survey. Seventy eight percent of respondents were at the

level of CIO or equivalent.

Summary. This study explores how IT entities cope with the proliferation of data found in

colleges and universities. Institutional data management challenges can be understood through

reflection of three broad domains of data impact: (a) the difficulties higher education institutions

face when attempting to retrieve, manipulate and analyze aggregate data for metrics and

planning; (b) the enormous body of content, primarily unstructured data. Unstructured data refers

to free form information outside the content restrictive environment of modeled fields of data.

Technology to compile, manipulate, and analyze unstructured data is considerably less mature

than what is available for structured data; and (c) the last domain that represents a data challenge

for college and universities is research data which exists in massive quantities. Digital data is

highly portable, sharable and searchable; qualities that some believe are leading us into an open-

access environment. However, research data in particular, hold unique problems related to

preservation, interpretation and ownership. Table 2 shows the level of confidence that higher

education administrators have in their institutions main information systems and data stores. The

mean response is based on a scale of 1 for strongly disagree, to 5 meaning they strongly agree.

The responses were primarily neutral to below neutral demonstrating a lack of conviction about

the institution's ability to ensure quality information in their key administrative systems. Solving

such problems could significantly bolster productivity. Failing to address these problems could

lead to lost data through deterioration or because no one is aware of its existence.

	Mean*	N	Std. Deviation
Characteristics of Major Data Elements			
A system of record is identified for each major data element.	3.54	301	0.981
Each major data element has a single definition that is recognized across the institution.	3.29	308	1.117
Each major data element is coded consistently across systems and data stores.	3.12	301	1.075
When the value of a major data element changes, the change propagates across all enterprise systems and data stores that use it.	3.11	308	1.140
When the value of a major data element changes, the change propagates across all business/academic unit ("shadow") systems and data stores that use it.	2.52	299	1.082
Data Quality Processes			
Processes are in place to assure data quality at the point of capture/origin (e.g., data entry).	3.24	306	0.975
Automated systems are in place to validate data across enterprise systems and data stores.	2.89	302	1.035
Processes are in place for documenting and reviewing all identified data quality issues.	2.76	304	0.995

*Scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Table 2. "Administrative Enterprise System Data Quality Measures" (p. 6).

Credibility. Yanosky holds a doctorate with the University of California at Berkeley. He is

currently employed with Richard N. Katz and Associates, a firm focused in institutional

effectiveness in higher education. In addition, he was the interim Director and research fellow at

EDUCASE Center for Applied Research where he authored several major studies on governance

in higher education. Earlier in his career Yanosky worked at Gartner, Inc. on their higher

education team.

Conclusions

This study describes how corporate-wide data governance practices in the private sector are used to enhance data quality **(**Olson, 2003; Panian, 2010; Redman, 2001, 2005), improve operational efficiencies (Otto, Wende, Schmidt, & Osl, 2007; Panian, 2010; Zornes, 2006), boost competitive advantage (Breur, 2009a**;** Lee, Pipino, Funk, & Wang, 2006; Wende & Otto, 2007; Whitehead, 2006), ensure regulatory compliance (Panian, 2010, Power & Trope, 2005), and instill confidence-based decision-making (Fisher, 2009; Redman, 2001; Sarsfield, 2009; Weller, 2008). The goal is to provide information that can pave the way for similar initiatives in institutions of public higher education. "Misunderstood and ignored, the extent of poor data quality continues to elude institutions," according to Whitehead (2006, p. 1). The literature suggests that poor data quality has negative ramifications across the entire organization (Breur, 2009a; Fisher, 2009). However, data governance practices are reported to maintain data of high-quality (Breur, 2009b; Lucas, 2010). The research reinforces the importance of holding the *organization* accountable for data governance and data quality and not placing the burden on *IT*, which is better suited to deal with IT governance from a technology viewpoint (Cheong & Chang, 2007; DAMA International, 2009; Wende, 2007).

The Root Causes of Poor Data Quality

The data quality quandary persists, according to McKnight (2009), because organizations don't comprehend the [operational] complexities that help to perpetuate data quality problems. The issue is complicated because data of poor quality "presents itself in so many different ways" (Mcknight, 2009, p. 32); the tremendous amounts of data being generated every day add further complexity. Higher education is a liberal contributor to the data boom by both consuming and generating large quantities of operational data (Yanosky, 2009). "Data is at once an

organization's greatest source of value and its greatest source of risk" (Soares, 2010). Table 3

provides an overview of the root causes of poor data quality.

Root Cause of Poor Data Quality	Citations
Aging data quality	McKnight, 2009; Sarsfield, 2009
Data anomalies	Sarsfield, 2009
Data entry quality	McKnight, 2009
Data processing quality	McKnight, 2009
Data silos	Sarsfield, 2009
Decaying data quality	Breur, 2009b
Degrading quality control processes	Olson, 2003
Delays in conveying information to decision makers	Olson, 2003
Disparate information systems	Fisher, 2009; Redman, 2008; Masayna, Koronios, & Gao, 2009
Duplications	Sarsfield, 2009; Weller, 2008
Inaccuracies	Olson, 2003; Redman, 2008; Sarsfield, 2009; Weller, 2008
Inadequate storage processes	Weller, 2008
Inadequately defined or coded data	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006; McKnight, 2009
Incomplete data	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006
Inconsistencies	Soares, 2010
Input RULES	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006
Integration issues	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006; McKnight, 2009;Olson, 2003
Internal competition	Sarsfield, 2009
Lack of standards	Sarsfield, 2009
Limited technology resources	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006
Lost data	Redman, 2008
Misinterpreted data	Lee, Pipino, Funk, & Wang, 2006; Redman, 2008
Multiple data sources	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006
Multiple Platforms [operating systems/different languages]	Lee, Pipino, Funk, & Wang, 2006; Sarsfield, 2009
Operational processes	Olson, 2003; Weller, 2008
Organizational confusion such as where is this data stored	Redman, 2008
Poorly articulated requirements	Olson, 2003
Security and accessibility	Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006; Redman, 2008; Weller, 2008
Unknown data	Sarsfield, 2009
Volume of data	Breur, 2009a; Fisher, 2009; Lee, Pipino, Funk, &

	Wang, 2006; Olson, 2003; Redman, 2001; Redman, 2008; Weller, 2008

Table 3. Summary of the root causes of poor data quality

The Importance of Addressing Data Quality Issues

Poor data quality can have a devastating impact on *costs* (Fisher, 2009; Olson, 2003;

Redman, 2001), due to things like rework expenses (Lucas, 2010; Olson, 2003), the cost to

implement new systems (Olson, 2003), the cost of time, revenue, and reputation (Redman,

2001), and lost opportunity (Lucas, 2010). In addition, the *believability factor* comes into play

when data is perceived as *so bad* that people no longer trust in it (Fisher, 2009; Lee, Pipino,

Funk, & Wang, 2006; Redman, 2001; Sarsfield, 2009) which can lead to the loss of production

and/or customers (Olson, 2003). The impact on *decision-making* can be equally profound when

based on anything less than data of high-quality (Redman, 2001; Weller, 2008), and the result

can be found to deplete morale (Redman, 2001). Expenses may translate to fines caused by *non-*

compliance of federal regulations (Fisher, 2009; Weller, 2008) and/or lead to *personal liability*

(Power & Trope, 2005; Redman, 2001).

Key Benefits of Data Quality Improvement and Chief Attributes of High Data Quality

Improved data quality can have a direct positive effect on risk mitigation (Fisher, 2009;

Weller, 2008) such as reductions in litigation expenses (Weller, 2008). When data is high

quality, organizations are able to better understand information sources and volumes of data are

more manageable (Weller, 2008). Security safeguards, trust levels, and customer satisfaction

improve; as well as revenue generation, and reduction in costs, and risks (Soares, 2010). Data of

high quality are accurate (Cheong & Chang, 2007; Olson, 2003; Panian, 2010), relevant,

complete, timely, and believable (Cheong & Chang, 2007; Olson, 2003). High quality data is

understood (Olson, 2003), accessible, auditable, consistent and secure (Panian, 2010) and conforms to business rules (Breur, 2009b).

Factors that Drive Data Quality and Improvement Principles

Federal regulations such as HIPAA, FERPA, Basel II and the Sarbanes-Oxley Act have helped to raise corporate awareness around the issue of data quality (Khatri & Brown, 2010). In order to avoid heavy non-compliance fines, organizations are paying closer attention to their data handling practices and internal control processes (Wallace, Lin, & Cefaratti, 2011). The literature shows a clear trend that more organizations are beginning to think strategically about their data in a number of ways: (a) as an information asset (Albrecht & Pirani, 2009**;** Breur, 2009a; Redman, 2005; Welch, 2009); (b) as an enterprise asset (DAMA International, 2009; Fisher, 2009; Gluck, 2008; Panian, 2010; Soares, 2011); (c) as a valuable asset (Fisher, 2009; Leek Pipino, Funk, & Wang, 2006) and (d) as a business asset (Olson, 2003; Redman, 2008). The desire to optimize revenue is another factor driving the interest in the assurance of high-quality data (Fisher, 2009; Lee, Pipino, Funk, & Wang, 2006; Panian, 2010) along with improving ROI (Breur, 2009a). Ultimately, when enterprise data are treated as assets there's an opportunity for organizations to improve upon their competitive advantage (Breur, 2009a; Les, Pipino, Funk, & Wang, 2006; Masayna, Koronios, & Gao, 2009; Redman, 2001). For example, when creatively integrating disparate information systems to improve data quality (Breur, 2009a) or when companies secure their proprietary data, organizations are likely to gain competitive advantage (Power & Trope, 2005).

Key Data Governance Tools and Practices

   Data governance frameworks. Data governance relies on tools including frameworks to

break down governance processes into manageable pieces, providing a structured approach to

data oversight (Berson & Dubov, 2011). Although there are various frameworks from which to

chose, there are also common elements among them. Table 4 shows the key elements found in

six selected frameworks. Some of the more prominent elements among frameworks include (a)

developing a data governance strategy (Masayan, Koronios, & Gao, 2009) or roadmap (Soares,

2011), (b) designing the data governance organizational structure to include boards and

committees, (c) developing data governance policies at the executive level, and (d) the

appointment of business data steward in particular (Berson & Dubov, 2011).

Data Governance Framework	Key Elements
Mike 2.0 Framework Berson & Dubov, 2011	Develop data governance strategy Design data governance organization Develop data governance policies Develop operational processes Perform data investigations and monitor specific data issues Implement data governance technology Design data governance architecture
Data Governance Institute Framework Berson & Dubov, 2011	Data quality management Design data governance architecture Develop data governance strategy Develop data governance policies Establish standards key emphasis on business process reengineering BPR, and enterprise data management, business intelligence and data warehousing Privacy, compliance and security
IBM Data Governance Council Framework & Maturity Model Berson & Dubov,	Also contains the maturity model used to assess at what stage of maturity a business is in Audit information Data quality management Design data governance architecture Design data governance organization Develop data governance policies

2011	Information lifecycle management data quality Metadata data quality Privacy, compliance and security Stewardship Value generation
Modified Data Governance framework Khatri & Brown, 2010	Data access policies Data lifecycle Data governance standards on data quality such as accuracy, credibility, completeness, and timeliness Data principles on intended use, metadata, and how data is accessed and by whom Metadata Physical metadata User metadata such as annotations on user preferences
Generic Stewardship Model Welch, 2009	Develop a data governance Council Domain Stewards, working directly below the governance council bridging the gap between IT and business Fosters alliance between IT and business Getting executive buy-in is critical Information stewards represent all others in the organization
Agile Data Governance Berson & Dubov, 2011	Chooses project implementation team Determines the size and extent of the data issue Recruits the data steward group

Table 4. Selected data governance frameworks and key elements.

Key roles and relationships. Weber, Cheong, Otto, and Chang (2008) identify five primary

roles in the data governance architecture: executive sponsor, chief data steward, technical data

steward, the data quality board and the business data steward. The role of business data steward

is particularly noteworthy as "it launches and controls the execution of a data governance

program" (Marco, 2006b, p. 17). It is the formal way to manage data assets as a representative of

key data stakeholders (Welch, 2009). Data stewards should manage their data oversight

responsibilities from a cross-organizational perspective, having a good understanding of the organization's vision (Gluck, 2008; Wende, 2007). It is important that the data is managed by the business side of the house, working in collaboration with IT (Cheong & Chang, 2007; DAMA International, 2009; Gluck, 2008; Weber, Cheong, Otto, & Chang, 2008). Some organizations distinguish between the roles of data stewards (Villar, 2009) wherein *executive data stewards* participate as part of a council, *coordinating data stewards* provide business steward team representation, and *business data stewards* are subject matter experts (DAMA International, 2009). Stewardship comes in different forms but many include a hierarchy of leadership on a data governance council or board who convene regularly to discuss data-related initiatives, policies, or problems (Berson & Dubov, 2001). The council provides senior-level, corporate-wide oversight (DAMA International, 2009; Redman, 2005; Welch, 2009). Some data stewardship steering committees are formed to provide support to the data governance council with assistance on relevant data related policies and assistance with stewardship team collaboration (DAMA International, 2009).

Data Quality and Data Governance in Higher Education

Colleges and universities are "drowning" in data continually growing at astounding rates (Yanosky, 2009) as a result of rapid advancements in technology (Albrecht & Pirani, 2009). Organizing data in meaning ways is essential for strategic planning and decision-making in institutions of higher education (Frost, Lucas, & Blankert, 2004). Sensitive, vital institutional data (Albrecht & Pirani, 2009) is more portable and can easily be copied or moved onto 32 gig portable thumb-size devices (Yanosky, 2009) [for example] in a matter of seconds.

All of this makes the management and security of highly sensitive data an increasingly intricate process, requiring oversight through data governance and stewardship, in which a wide-

range of custodial responsibilities are implemented to ensure commitment to secure, accurate, and readily accessible data. Stewardship involves the collaboration of units across the entire organization. As noted by Albrecht and Pirani (2009), "… to gain a complete perspective, one must ascertain the user's point of view" (p. 11). Data governance programs need to be an institutional endeavor administered centrally, and accepted throughout the college community (Albrecht & Pirani, 2009). A breach of data governance best practices was brought to light in Yanosky's (2009) study wherein "formally assigning responsibility for managing data resources to data stewards was the exception rather than the rule. Only a third of the institutions conveyed having a formal policy defining data steward responsibilities, though slightly more than half of institutions with FTE enrollments great than 15,000 had them" (p. 7).

   Drivers of change. There are a number of key factors that are driving the need to adopt data governance practices in higher education. These include: (a) the opportunity to investigate new revenue sources such as non-credit instruction, and (b) the need to gain competitive advantage in the face of declining enrollment (Drucker, 2005); (c) the need to mitigate risk given the institution's handling of extremely sensitive data, alleviating apprehension surrounding existing data handling practices, and (d) the need to create a more centralize data environment (Albrecht & Pirani, 2009). Additional drivers include (e) the need to ensure greater transparency and accountability, (Drucker, 2008), and (f) meet compliance mandates leads to a reassessment of information handling practices (Albrecht & Pirani, 2009; Drucker, 2008); and (g) strong support from senior administrator reinforced [data governance] activities (Albrecht & Pirani, 2009).

   Lessons learned.

Institutions of public higher education will benefit from improved data quality and data governance. Seventeen lessons learned are derived from selected institutions (see table 5). Leadership at the highest level must be visibly behind the push to instill data governance across the entire institution (Drucker, 2005). Expectations must be realistic and encourage a mission-critical attitude toward a data-driven culture (Albrecht & Pirani, 2009). Tools should be provided, and data handling operations should be reviewed periodically to ensure data is of the highest quality (Albrecht & Pirani, 2009). Communication must be clearly articulated mandating data governance as a business imperative, engaging stakeholders early on in a well-coordinated cross-organizational effort (Albrecht & Pirani, 2009). It is important not to place blame when approaching the topic of data governance but rather encourage a cooperative spirit based on a mutual desire to achieve a common objective (Albrecht & Pirani, 2009). Business rules must be easily comprehended and well-defined in order to ensure sound judgment from which to base decision-making (Albrecht & Pirani, 2009). Research data, a prominent component in higher education, has its own distinct issues in regard to interpretation, ownership and preservation (Yanosky, 2009). Time and resources should be devoted to workflow, management and architecture of institutional data (Drucker, 2005); it should also be viewed as an enterprise asset, highly valued, protected, and actively shared across divisional boundaries (Albrecht & Pirani, 2009).

Lessons Learned	Citations
Institutional leadership must become respectful of data and astute in using data to inform institutional decisions; institutions must devote time, effort, and resources, to information architecture, to workflow, and to data management.	Drucker, 2005, p. 6.

With a contemporary or well-architected enterprise systems environment, broadly understood data models, clear business rules, and reasonably clean data, the promise of the future academic analytics environment is within reach.	Drucker, 2005, p. 7.
Senior administration is mission critical.	Albrecht & Pirani, 2009, p. 19.
When approaching business units about data handling practices it's important not to place blame but rather in the spirit of making things better; this approach fosters a more cooperative attitude.	Albrecht & Pirani, 2009, p. 19
Research data has unique issues surrounding interpretation, ownership and preservation.	Yanosky, 2009
Unstructured data is more difficult to manage than structured data.	Yanosky, 2009
Set realistic expectations.	Albrecht & Pirani, 2009
Plan and implement for the future—it's a culture shift that doesn't happen overnight.	Albrecht & Pirani, 2009
Maintain a central focus. "The more you convey a well-coordinated, central focus and an overall institutional mandate, the better off you will be."	Albrecht & Pirani, 2009, p. 20
Engage stakeholders early on in the process.	Albrecht & Pirani, 2009
Clear communication and governance enhance user guidance and structure.	Albrecht & Pirani, 2009; Drucker, 2005
Provide tools and detailed processes to facilitate compliance.	Albrecht & Pirani, 2009
Acknowledge institutional data as an important institutional asset.	Albrecht & Pirani, 2009
Identify new streams of revenue.	Drucker, 2005
Improve [data handling practices] that impede decision-making.	Drucker, 2005
Work to shift the culture from "just-in-time"	Ducker, 2005

delivery to predictive.	
Shift culture from reactive to proactive.	Albrecht & Pirani, 2009

Table 5. Lessons learned in selected institutions of public higher education.

References

Albrecht, B. & Pirani, J.A. (2009). Revitalizing data stewardship through risk reduction:

Managing sensitive data at the University of Virginia**,** *EDUCAUSE Center for Applied*

Research, Case study 8, retrieved from http://www.educause.edu/ecar

Badrakhan, B. (2010, Aug). Drive toward data governance. *Electrical Wholesaling*, *91*(8), 35-48.

Retrieved from

http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr

ue&db=buh&AN=53157211&site=ehost-live

Berson, A., & Dubov, L. (2011). Master data management and data governance. *McGraw-Hill*

Osborne Media; 2 edition.

Breur, T. (2009a). Data quality is everyone's business--designing quality into your data

warehouse--Part 1. *Journal of Direct, Data and Digital Marketing Practice, 11* (1)

doi:10.1057/dddmp.2009.14. Retrieved from

http://find.galegroup.com.ezmw.ez.cwmars.org:4200/gtx/infomark.do?&contentSet=IAC-

Documents&type=retrieve&tabID=T002&prodId=AONE&docId=A205907920&source=gal

e&srcprod=AONE&userGroupName=mlin_c_wachcc&version=1.0

Breur, T. (2009b). Data quality is everyone's business – managing information quality—Part 2.

Journal of Direct, Data and Digital Marketing Practice, *11* (2) 114-123. Retrieved from

http://www.palgrave-journals.com/dddmp/journal/v11/n2/pdf/dddmp200921a.pdf

Busch, C., De Maret, P.S., Flynn, T., Kellum, R., Le, S., Meyers, B., White, R., & Palmquist, M.

(2005). *Writing@CSU. Colorado State University Department of English*. Retrieved from

http://writing.colostate.edu/guides/research/content/com3b3.cfm

Buzydlowski, J., Hand, J., Song, L., & Hassel, L. (n.d.) Statistical techniques to facilitate the

population of a data warehouse. *College of Information Science and Technology, Drexel*

University. Retrieved from

http://www.ischool.drexel.edu/faculty/song/publications/p_CSI98-DW.pdf

Chapman, A. D. (2005). *Principles of Data Quality* (version 1.0.). Report for the Global

Biodiversity Information Facility, Copenhagen. Retrieved from

http://www2.gbif.org/DataQuality.pdf

Cheong, L. K., & Chang, V. (2007). *The need for data governance: A case study. 18th*

Australasian Conference on Information System. Paper presented at the 18th Australasian

Conference on Information System, Toowoomba, Australia. Abstract retrieved from

http://www.acis2007;usq.edu.au/assets/papers/171.pdf

DAMA International. (2009). *The DAMA guide to the data management body of knowledge* (1st

ed.). Bradley Beach, NJ 07720. Abstract retrieved from

www.technicspub.com/books.htm.

Drucker, P. (2005). Academic analytics in the future of higher education. *EDUCAUSE Center*

for Applied Research. Retrieved from

http://net.educause.edu/ir/library/pdf/ers0508/rs/ers05089.pdf

Feldman, S. (2004). The high cost of not finding information. *KM World.* Retrieved from

http://www.kmworld.com/Articles/PrintArticle.aspx?ArticleID=9534

Fisher, T. (2009). The data asset: How smart companies govern their data for business success.

John Wiley & Sons, Inc., Hoboken, N.J.

Frost, J., Lucas, C., & Blankert, J. (2004). Strategies and tools used to collect and report strategic

plan data, Online Submission. Paper presented at the *Annual Forum of the Association for*

Institutional Research (AIR) (44th, Boston, MA). Retrieved from

http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED491018

Gantz, J., & Reinsel, D. (2010, May). The digital universe decade: Are you ready? *IDC*

International Data Corporation. Retrieved from http://idcdocserv.com/925

Gibbs, M., Shanks, G., & Lederman, R. (2005). Data quality, database fragmentation and

information privacy. *Surveillance & Society, 3*(1):45-58. Retrieved from

http://www.surveillance-and-society.org/Articles3(1)/data.pdf

Gluck, K. (2008). Information stewardship: Roles and responsibilities in Vattenfall's

maintenance process. *Masters' Thesis, ETH Electrical Engineering.* Retrieved from

http://www.isaca.se/dynamaster/file_archive/091216/3d6026d5e367614f125b0bb3d127dec8/

Uppsats.pdf

Hoover, S. & Schubert, N. (n.d.). 9 KPIs successful construction firms should monitor.

Construction: Effective Management of Projects for General Contractors and Homebuilders.

Retrieved from http://www.microsoft.com/en-us/dynamics/industries/construction-erp.aspx

Jennings, M. (2004, August). Business rules, business meta data and data stewardship

component of the complete meta data model. Enterprise Architecture View*, DM Review, 14*

(9), 79-81. Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=145

32955&site=ehost-live&scope=site

Juran, J. M. (July-August 1993). Made in U.S.A.: A renaissance in quality. *Harvard Business*

Review, retrieved from http://hbr.org/product/made-in-u-s-a-a-renaissance-in-

quality/an/93404-PDF-

ENG?Ntt=made%2520in%2520u.s.a.%253A%2520%2520renaissance%2520in%2520qualit
y

Khatri, V., & Brown, C., V. (2010). Designing data governance. *Communications of the ACM*,

   53(1), 148-152. Retrieved from

   http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr
   ue&db=buh&AN=47220627&login.asp?custid=wchstcc&site=ehost-live

Lee, Y. W., Pipino, L. L., Funk, J. D., & Wang, R.Y. (2006). Journey to data quality.

   Massachusetts Institute of Technology (MIT) Press, Cambridge, MA.

Legacy systems stifle business options. (2008) *Telecom Asia 19* (24-25) 2p Retrieved from

   http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=cph&AN=313
   75270&login.asp&site=ehost-live&scope=site

Levi, Y. (2008). Digital preservation: An ever-growing challenge. *Information Today*, *25*(8), 22.

   Retrieved from

   http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr
   ue&db=buh&AN=34459911&login.asp?custid=wchstcc&site=ehost-live

Lucas, A. (2010). Corporate data quality management in context. *LaboratorioNacional de

   Engenharia Civil and Instituto Superior de Ecohomia e Gastaok Portugal (research in

   progress).* Retrieved from http://hdl.handle.net/10400.5/2271.

Marco, D. (2006a). Understanding data governance and stewardship, Part 1. *DM Review*.

   Retrieved from

   http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=225
   46843&site=ehost-live

Marco, D. (2006b). Understanding data governance and stewardship, Part 2. *DM Review*.

Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=235
07978&site=ehost-live

Marinos, G. (2004, August). Data quality: A risk-based approach. *DM Review, 14* (8) 10-14.

Retrieved from

http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr
ue&db=buh&AN=13996993&login.asp?custid=wchstcc&site=ehost-live

Masayna, V., Koronios, A., & Gao, J. (2009). A framework for the development of the business

case for the introduction of data quality program linked to corporate KPIs & governance.

2009 Fourth International Conference on Cooperation and Promotion of Information

Resources in Science and Technology, pp.230-235. Retrieved from

http://doi.ieeecomputersociety.org/10.1109/COINFO.2009.56.

McGilvray, D. (2006). Data governance: The missing link for data quality success. *San*

Francisco Bay Area, Data Management Association presentation for DAMA Day 2007.

Retrieved from

http://sfdama.org/Presentations/2006/dama_day_2006/McGilvray_Data_Governance.pdf

McKnight, W. (2009). Seven sources of poor data quality. *Information Management (1521-*

2912), 19 (3), 32-33. Retrieved from EBSCOhost

http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr
ue&db=buh&AN=43862780&site=ehost-live

Merriam-Webster Dictionary (2011). Stewardship. Retrieved from http://www.merriam-

webster.com/dictionary/stewardship

Olson, J. (2003). *Data quality: The accuracy dimension.* San Francisco, CA: Morgan Kaufmann

    Publishers.

Otto, B., Wende, K., Schmidt, A., & Osl, P. (2007, Dec). *Towards a Framework for Corporate*

    Data Quality Management. Paper presented at the 18th Australasian Conference on

    Information System, Toowoomba, Australia. Retrieved from

    http://aisel.aisnet.org/acis2007/109

Panian, Z. (2010). Some practical experiences in data governance. *World Academy of Science*

    Engineering and Technology. Summary retrieved from

    http://www.waset.ac.nz/journals/waset/v62/v62-174.pdf

Petschulat, S. (2010, January). Other people's data. *Communication of the ACM, 53* (1), 53-57.

    Retrieved from

    http://ezmw.ez.cwmars.org:4200/login?url=http://search.ebscohost.com/login.aspx?direct=tr

    ue&db=buh&AN=47220612&login.asp?custid=wchstcc&site=ehost-live

Power, M. E., & Trope, R. L. (2005). Sailing in dangerous waters: A director's guide to data

    governance. *American Bar Association, Chicago, IL 60610.*

Redman, T. C. (1998). The impact of poor data quality on the typical enterprise: Poor data

    quality has far-reaching effects and consequences. *Communications of the ACM, 41*, No. 2.

    doi: 10.1145/269012.269025. Retrieved from

    http://delivery.acm.org.libproxy.uoregon.edu/10.1145/270000/269025/p79-

    redman.pdf?ip=128.223.86.31&CFID=26516908&CFTOKEN=77330878&__acm__=13079

    91967_cf2f48469f2391fdc16e290d26608919

Redman, T. C. (2001). Data quality: The field guide. *Butterworth-Heinemann*, Woburn, MA

    01801-2041.

Redman, T. C. (Aug, 2004). Data: An unfolding quality disaster. *DM Review*, *14*(8), 21-23.

Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=139

97035&site=ehost-live

Redman, T. C. (Sept/Oct 2004). Data quality: Should universities worry? *EDUCAUSE Review,*

39, No. 5:12-13. Retrieved from http://net.educause.edu/ir/library/pdf/ERM0457.pdf

Redman, T. C. (2005). A comprehensive approach to data quality governance. *A Navesink*

Consulting Group White Paper. Retrieved from http://iaidq.org/webinars/doc/2006-11-

redman-data-governance-white-paper-2005-06-22-version.pdf

Redman, T.C. (2008). Data driven: Profiting from your most important business asset. Retrieved

from: *Harvard Business School Publishing*, Boston, MA 02163.

Russom, P. (2006). Taking data quality to the enterprise through data governance. *The Data*

Warehouse Institute, A 101 Communications Publications. Retrieved from

ftp://ftp.software.ibm.com/software/uk/govern/data_quality_to_the_enterprise.pdf

Sarsfield, S. (2009). *The data governance imperative. A business strategy for corporate data.*

Cambridgeshire, United Kingdom: IT Governance Publishing.

Schutzer, D. (2010). Risk management and data governance. *The CTO Corner, Financial*

Services Roundtable Technology Group. Retrieved from

http://www.bits.org/downloads/Publications%20Page/CTOCornerOctober2010.pdf

Soares, S. (2011). The IBM data governance unified process: Driving business value with IBM

software and best practices. *MC Press Online*, LLC, Ketchum, ID.

Teodorescu, D. & Frost, S. (2002). IR professionals in the market of institutional knowledge:

Envisioning new roles for our work. Paper presented at the *42nd Association for Institutional*

Research Forum, Toronto, Canada.

Thomas, G. (2006). Alpha males and data disasters: The case for data governance. *Brass Cannon*

Press, Orlando, FL. Retrieved from

http://www.datagovernance.com/part_1of3_datagovebook.pdf

Thomas, G. (2009). How to use the DGI data governance framework to configure your program.

The Data Governance Institute. Retrieved from

http://www.datagovernance.com/wp_how_to_use_the_dgi_data_governance_framework.pdf

Villar, M. (2009). Establishing effective business data stewards. *Business Intelligence Journal*,

14 (2) 23-29. Retrieved from

http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=415

22172&site=ehost-live

Vosburg, J. & Kumar, A. (2001). Managing dirty data in organizations using ERP: Lessons from

a case study. *Industrial Management & Data Systems, 101* (1), p.p. 21-31. Retrieved from

http://www.emeraldinsight.com/journals.htm?issn=0263-

5577&volume=101&issue=1&articleid=850013&show=abstract

Wallace, L., Lin, H., & Cefaratti, M. (2011). Information security and Sarbanes-Oxley

compliance: An exploratory study. *Journal of Information Systems*, *25* (1), 185-211. Abstract

retrieved from

http://content.epnet.com/pdf25_26/pdf/2011/JIN/01Mar11/59414945.pdf?T=P&P=AN&K=5

9414945&EbscoContent=dGJyMNLr40SeprM4v%2BbwOLCmr0meqK5Ssaa4Sq%2BWxW

XS&ContentCustomer=dGJyMPGuslGwrbVNuePfgeyx%2BEu3q64A&D=aph

Weber, K., Cheong, Otto, B., & Chang, V. (2008). Organizing accountabilities for data quality

   management – A data governance case study. *Institute of Information Management,*

   University of St. Gallen, Muller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland. Retrieved

   from

   http://web.iwi.unisg.ch/org/iwi/iwi_pub.nsf/wwwPublYearGer/028D5245B5975BF7C12574

   F1002EBB0E/$file/DW2008%20FP%20Data%20Governance%2001%20kwe.pdf

Weill, P. (2004). Don't just lead, govern: How top-performing firms govern IT. *Mis Quarterly*

   *Executive, 3 (1). R*etrieved from http://www.umsl.edu/~lacitym/topperform.pdf

Welch, M. (2009, June). *Information governance and stewardship for records and information*

   management. Abstract retrieved from

   https://www.cis.unisa.edu.au/wki/images/5/58/Welch_thesis_June_01_2009.doc.

Weller, A. (2008). Data governance: Supporting datacentric risk management. *Journal of*

   Securities Operations & Custody, 1(3). Retrieved from

   http://search.ebscohost.com.libproxy.uoregon.edu/login.aspx?direct=true&db=bth&AN=339

   38168&site=ehost-live&scope=site

Wende, K., Otto, B. (2007). A contingency approach to data governance. *University of St.*

   Gallen, Switzerland. Retrieved from

   http://mitiq.mit.edu/iciq/Documents/IQ%20Conference%202007/Papers/A%20CONTINGE

   NCY%20APPROACH%20TO%20DATA%20GOVERNANCE.pdf

Wende, K. (2007). *A model for data governance – Organizing accountabilities for data quality*

   management. Paper presented at the 18th Australasian Conference on Information System,

   Toowoomba, Australia. Retrieved from http://aisel.aisnet.org/acis2007/80.

Whitehead, J.C. (2006). Solving the data quality problem. *TriTuns Innovation, LLC*. Retrieved

from

http://www.itpoint.be/uploads/media/Solving_The_Data_Quality_Problem__Trituns_Innovat

ion.pdf:

Whiting, R. (2006, May, 8). Hamstrung by defective data: Business information that's redundant,

outdated, or flat-out wrong trips up organizations large and small—but there are fixes in the

offing. *InformationWeek*. Retrieved from

http://www.informationweek.com/news/187200771

Wolter, R. (2007, January, 18). MDM master data management hub architecture – Maintenance

and operation [Web log post]. Retrieved from

http://blogs.msdn.com/b/rogerwolterblog/archive/2007/01/18/mdm-master-data-

management-hub-architecture-maintenance-and-operation.aspx

Yanosky, R. (2009). Key findings: Institutional data management in higher education. *Educause*.

Retrieved from http://net.educause.edu/ir/library/pdf/EKF/EKF0908.pdf

Zornes, A. (2006). Corporate data governance best practices. *The CDI Institute MarketPlus TM

In-Depth Report.* Retrieved from

http://www.tcdii.com/PDF/Data_Governance_white_paper_-_April_2006.pdf