Presented to the Interdisciplinary Studies Program:

UNIVERSITY OF OREGON
APPLIED INFORMATION MANAGEMENT

Applied Information Management
and the Graduate School of the
University of Oregon
in partial fulfillment of the
requirement for the degree of
Master of Science

# Balancing Data Security and Accessibility in Higher Education

CAPSTONE REPORT

**Stephen J. Brouwers**
**Applied Information Technology Specialist**
**Clackamas Community College**

University of Oregon
Applied Information
Management
Program

**May 2016**

Academic Extension
1277 University of Oregon
Eugene, OR 97403-1277
(800) 824-2714

Approved by

_____

Dr. Kara McFall
Director, AIM Program

# Balancing Data Security and Accessibility in Higher Education

Stephen J. Brouwers

Clackamas Community College

**Abstract**

Regulatory requirements and the fact that data breaches are on the rise both beg for increased data security, while ever-growing expectations within higher education to be more efficient and help students succeed demands data be more accessible to answer endless business questions. It is now more essential than ever that higher education institutions determine how to balance data security with accessibility – two major aspects addressed by data governance. This literature review identifies tools and techniques that can help leadership within higher education institutions balance data security with the high levels of accessibility necessary to provide value to the organization and answer key business questions in a timely manner.

*Keywords:* accessibility, balance, data, governance, higher education, policy, security, stewards, stewardship

# Table of Contents

## Contents

## List of Tables and Figures

## Introduction to the Annotated Bibliography

### Problem

Data breaches are on the rise, with nearly 20% of the publicly disclosed breaches between 2005 and 2012 impacting the education sector (Sen & Borle, 2015). Institutions of higher education in the United States are also bound by the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99), which is a Federal law protecting the security and privacy of student education records that places the burden of protecting student data squarely on the campuses (Family, n.d.). These statistics and regulations have brought data security to the forefront of conversation within the higher education sector in recent times.

At the same time expectations that require highly accessible data have grown significantly. Many states are changing public funding models from simple enrollment-based models to those that are based on student success (Performance, 2015), requiring an ongoing commitment to institutional research, and a need to use data to help find ways to predict student problems and find ways to help students succeed. In relation to this, states are also beginning to implement regulations such as Oregon's 40-40-20 mandate that dictates 40% of Oregon citizens should earn a 4 year degree, 40% should earn a 2 year degree or other post-high school certificate, and the remaining 20% should earn a high school diploma or equivalent by 2025 (Hamilton, 2014).

Petersen (2012) notes that "data helps unlock the mystery about college costs, learning outcomes, institutional effectiveness, and other performance indicators" (p. 46) that all relate to student success. Unfortunately many organizations find themselves locked into outdated models where IT strictly controls access to data and uses the language of security to frame an approach of exclusion and denial (Swoyer, 2016). Ferrel and Hartline (2011) state that while an

organization's own records are often the best source of data to help achieve an organization's goals; one of the biggest problems with internal data is that data are often not in a readily accessible form. This leads to a definition of accessibility that not only includes adequate security rights to access data, but ensuring that data are in an efficiently usable format, and that there are adequate tools and expertise to turn data into actionable information. The problem is that with all the legitimate needs for securing data, higher education institutions also need to determine how they can ensure an adequate level of accessibility so that the data can be used to help meet strategic objectives.

**Purpose**

Watson and McGivern (2016) define data governance as "the people, processes, and technologies used to manage, protect, and use data so that organizations can leverage it as an organizational asset with the ultimate goal of creating value for the organization" (p. 5). The purpose of this literature review is to explore key elements of data governance in order to provide proven techniques or methodologies that a higher education institution can utilize to balance the level of accessibility and use of data necessary to meet strategic objectives with an appropriate level of security.

**Research Question**

To support this purpose, the overriding question is as follows:

What are the best practices for balancing data security and data accessibility through governance to ensure the means to create strategic business value while minimizing risk?

The following question acts to guide the structure of this study:

- On which specific elements of data governance should a higher education institution focus in order to achieve this balance?

**Audience**

Chapple (2013) states "it is rare to find an effective data governance program that began as a grassroots effort" (p. 20). It is well documented that executive support is critical to the successful implementation of data and information governance models (Bowen & Smith, 2014; Chapple, 2013; Kooper et al., 2011; Weber et al., 2009). This is partly the case because there are several obstacles in getting many areas of the business on the same page (Mont, 2015). That being the case, the primary audience for this study is executive leadership that includes the Chief Financial Officer (CFO), Vice President (VP) of Instruction and Student Services, Chief Information Officer (CIO), deans and associate deans within the leadership cabinet, as well as the Registrar and Director of Institutional Research. These particular executives and high-level managers are key stakeholders for this study because they not only have the decision and policy making authority that is required to create and enforce a culture and environment for appropriate data security and accessibility; they also have a vested interest as key consumers of data for other decision-making needs. Hickson and Dowdy (2014) state that executive support is necessary to provide the leadership and resources necessary to unite an institution through strategic vision and procedural expectations.

In no small way however, all employees are stakeholders in governance (Kooper et al., 2011), and there is certainly a wider audience that consists of a cross-functional array of representatives from each major area of campus, as well as the Enterprise Applications Team within IT that generally works at the development level creating and maintaining the access structures and organizing base datasets from which all of this is based.

**Search Report**

Data and Information Governance are fairly large topics for which there is a lot of information. Starting with a general UO Library Article Search located at https://library.uoregon.edu/ a combination of search terms were used as described below. The Business Source Complete database is a major source that yields a lot of articles on data governance and data security in general, but the EduCause site contains some higher education specific articles.

Reference lists from vetted articles were followed in order to find new articles. Google Scholar and Google proved to be useful in order to refine the detailed titles and author names found within other article's reference lists. Other potential databases in the UO library have not been found to yield many relevant results.

**Reference evaluation criteria.** References are evaluated based on criteria as defined by the Center for Public Issues in Education that includes authority, timeliness, quality, relevancy, and bias (Evaluating, 2014). A reference is deemed authoritative if it is published in a peer-reviewed journal and the author has affiliation with a reputable organization and/or related higher education credentials. Due to the fast-changing pace in the world of data governance and security, references are considered timely if they were produced within the past 5 years. Quality is reviewed by examining the grammar, spelling and overall structure of the document. References are determined relevant to the topic, which means that they have to do with approaches or tools that address data accessibility, security or the relationship between the two – preferably within a higher education environment, but not necessarily. All references are also examined to determine if they express any particular bias that may indicate they are more of a sales pitch or propaganda tool than true research or other credible information. A few exceptions are given to the timeframe if they are academic works that otherwise fit the specified criteria but

were either out of the date range or not peer-reviewed in cases where the content was specific to the study.

**Documentation approach.** References are tracked using two methods. The standalone version of Zotero houses the basic bibliographic information and a Word document is used to keep a running annotated bibliography and any other associated notes about the reference. While Zotero has proven a handy resource to export citations from the University of Oregon Library and has capabilities beyond its use for this research, it has been found to not always produce accurate APA formatted reference listings, and as a new tool to the author cannot be fully trusted as a sole resource to track the references.

**Key terms.**

- data governance
- information governance
- governance
- data security
- data breach

**Sub-terms.**

- higher education
- education
- strategy
- model
- strategic priorities
- objectives
- accessibility
- stewards
- stewardship
- policy
- procedure

## Annotated Bibliography

The following Annotated Bibliography includes references that examine key aspects of data governance related to data security or data accessibility. References are selected that help point out key concerns and solutions in regard to these areas and aid in determining best practices for balancing them, particularly in a higher education environment.

Each annotation includes three elements: (a) a full bibliographic citation; (b) an abstract from the author when available; and (c) a summary detailing its relevance to this research. All ideas included in an annotation are representative of ideas presented by the author(s) of the reference. Paraphrased contents are not cited, but direct quotes contain in-text citations indicating the page number of the reference.

### Data Governance in Education

Chapple, M. (2013). Speaking the same language: Building a data governance program for institutional impact. *EDUCAUSE Review*, *48*(6), 14–16.

**Abstract.** Colleges and universities should be among the world's leading institutions in the field of data governance. After all, higher education institutions are dedicated to the creation and dissemination of knowledge. Why, then, do those who work in colleges and universities often have so much difficulty corralling information about their own operations and using it to share a consistent story with their stakeholders? This article describes the five pillars of Notre Dame's Data Governance Network, whose design emphasizes two very important points about data governance. First, placing "Access to Data" at the top of the model communicates a clear end-goal of the program: providing individuals who have legitimate business needs with the ability to access the data they need in a timely, effective manner. Second, placing "Technology" at the base of the

model conveys that data governance programs are not all about technology. Although technology may serve as a foundational tool for the development of strong data practices, these remain business processes that are "supported by" technology. The platform that supports data-driven decision-making across the institution should build toward the common goals of creating a data environment that embraces the five pillars of Quality & Consistency, Policies & Standards, Security & Privacy, Compliance, and Retention & Archiving.

**Summary.** Chapple states that data governance is used to help organizations effectively share information across functional units and bridge gaps between information systems. He uses Notre Dame's data governance framework as a primary example, which that places "Technology" at the base to represent a foundation and "Access to Data" at the top to signify that it is the end goal - including 5 pillars that are considered to be the disciplines that allow users to leverage technology to gain adequate access to business data.

Chapple places an emphasis on the "Security & Privacy" pillar. He states that as far as regulations, FERPA is the main one to consider when examining security and privacy in the Education sector, but also notes that there are other regulations to consider that include The Health Insurance Portability and Accountability Act (HIPPA), The Gramm-Leach-Bliley Act (GLBA), and The Payment Card Industry Data Security Standard (PCI DSS). In reference to implementation of security to achieve the end goal of adequate and appropriate access, Chapple discusses a couple of different approaches. The approach of utilizing IT and the CIO as the center of responsibility for data governance activities is referenced as the traditional approach, but states how other

approaches such as utilizing the Institutional Research (IR) office to lead these efforts are successful because of their close tie to the business end of the college and their knowledge and use of the data. University of Nevada, Las Vegas (UNLV) is mentioned as an example institution that converted their IR office to an Office of Decision Support and gave it all the related responsibilities and states that the key characteristics for the person in charge of data governance is to have a diverse background, the ability to build good relationships, and a good understanding of the institutions data and its business processes.

In order to "speak the same language", Chapple states that creating a data dictionary that provides shared definitions is essential. He introduces the Responsible-Accountable-Consulted-Informed (RACI) matrix as the method of choice. This matrix not only defines the terms, but also makes the roles of each stakeholder official and transparent for each term in the data dictionary. This article's relevance to this paper is based on the fact that it makes obvious that defining data and defining explicit roles for each data term's stakeholder in this manner, security is much more obvious, which in turn allows institutions to the ensure data is made accessible to the appropriate stakeholders and not to others.

Hickson, J., & Dowdy, J. G. (2014). Organizational structures for data governance at community colleges. *Education Advisory Board: Community College Forum*.

**Abstract.** Institutions already must comply with data reporting regulations and increasingly desire to leverage data as a predictive tool to improve academic and business outcomes. This report focuses on data governance at community colleges and draws upon promising practices from four-year institutions and the private sector to analyze

personnel needs and strategies for developing a data governance policy. Key observations from our research include: (a) data governance policies require support from executive leadership; (b) successful data governance policies include centralized documentation and a data dictionary; (c) administrators maintain strategic vision for data governance and generate policies and procedures through the delegation of small projects; and (d) college personnel establish data governance councils and appoint data stewards.

**Summary.** Hickson and Dowdy is relevant to this research because it clearly points out that the data security and access needs of an institution can be realized and achieved by providing clear data ownership and data steward roles, defining and categorizing data, and creating a data council that has distinct task-based workgroups. They give several charts and tables that can be used to develop and document those roles and data definitions, as well as guidance on how to perform these tasks.

Kelly, M. (2015, June 8). The chief data officer in higher education. *EDUCAUSE Review*. Retrieved from: http://er.educause.edu/articles/2015/6/the-chief-data-officer-in-higher-education

**Abstract.** Kelly describes the University of North Carolina's path to collaborative data management and the rise of his position of Chief Data Officer (CDO). The article explains the necessity of data stewardship, collaboration, and the fact that data roles are moving beyond matters of access and security and into a need for focused, active and intentional data management. There is further discussion of data standards, data quality and integrity assurance, but this articles primary focus is on the need for dedicated leadership along with shared ownership of data across the institution.

**Summary.** Kelly discusses the University of North Carolina's path to collaboratively manage data access. They find that their initial attempts to create a collaborative data environment, a Data Access Advisory Committee of data stewards that was supposed to make recommendations to the CIO and attempting to create data definitions and standards, is not enough without some additional dedicated management. A role of strategic information manager evolved into higher-level Chief Data Officer (CDO). The CDO is the subject matter expert on data governance principles, policy, and practices, and facilitates activities to address three compelling data challenges: compliance and reporting obligations, data risks, and data quality.

They quote Jane Griffin, a principal with Deloitte Consulting LLP, as saying that "effective CDOs are those individuals who possess a balance of technical skill, business knowledge, and people skills to smoothly navigate the technical and political hurdles of shepherding valuable corporate data" (para. 18). A CDO will also need strong and trusting relationships with other stakeholders, including institutional research and assessment experts, analytics teams embedded in functional and academic areas, the information security team, director of internal audit, risk manager, compliance officer, institutional review board director, general counsel, IT operations, and academic research leadership.

Kelly's article is relevant to this paper because it not only backs up the fact that definitions of data and roles helps determine adequate security, but adds that it takes dedicated leadership and a data council made up of the proper stakeholders to put this together.

Petersen, R. (2012). Policy dimensions of analytics in higher education. *EDUCAUSE*

*Review*, *47*(4), 44–46.

**Abstract.** The current higher education landscape is replete with demands for improving accountability, increasing efficiency, and controlling costs. At the same time, information technologies make it easier to collect and analyze information to measure outcomes or to assist in decision making. Consequently, there is a higher demand for better information and also a never-before-available supply of data. This corresponding increase in both supply and demand creates the perfect storm for higher education to move into a new era of analytics. However, in a recent discussion session about the legal issues and campus policy dimensions for analytics, one participant joked: "This is where analytics efforts on campus come to die!" Although legal and policy dimensions do present considerable challenges, they should not be simply dismissed as obstacles with no solutions. Therefore, it is helpful to frame the policy context of analytics to include the policy drivers that are generating demand and the campus policy choices that should be guiding campus practice.

**Summary.** Data from both internal and external sources help provide greater accountability and transparency within college operations and finance. Colleges continue to move away from when data was collected, analyzed and reported primarily to satisfy regulatory requirements and there is an ever-increasing recognition of the value of data-informed decisions and therefore a growing realization that data needs to be managed. This is relevant to this research because it gets at the need for data accessibility in this way and discusses that the main tasks to properly manage data include classifying and de-identifying data, defining roles and responsibilities, and creating data policies.

Classifying data into categories helps determine the necessary level of protection for each data element (such as public, non-public and sensitive), and de-identifying data reduces the risks associated with data – both aiding in security while giving higher education institutions the ability to increase access for business use. Explicitly spelling out roles and responsibilities of data stewards and data custodians allows institutions to be adequately selective in giving access so that appropriate access is granted on a need-to-know basis in order to meet business needs while smartly reducing access where it is not needed or useful. Related policies should then be developed to explicitly define access and authorization, ensuring security is transparent, explicit and understood, not viewed as a barrier.

**Data Governance - General**

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. http://doi.org/10.1145/1629175.1629210

**Abstract.** We have presented a data governance framework that can be used by practitioners to develop a data governance strategy and approach for managing data as an organizational asset. We have identified five decision domains, presented arguments for why each of these domains is important, described some key decisions to be made for each domain, and provided some examples of organizational positions that may be given accountability. We also have proposed that differing levels of centralized, decentralized, and shared decision rights may be appropriate for different decision domains in the same organization.

**Summary.** Khatri and Brown suggest that data governance refers to who holds the decision rights and is held accountable for an organization's decision-making about its

data. They propose a framework that contains five interrelated and overlapping decision domains that include: (a) data principles; (b) data quality; (c) metadata; (d) data access; and (e) data lifecycle. Data principles is placed at the top of the framework because it establishes the link to the business, where the business use of the data includes how data are interpreted and shared, as well as the regulatory environment around data use. Data access covers the security, risk, backup and access standards.

The authors state that there should be a clearly defined business owner of data and that the business owner must have an important role in managing access. This is very relevant to this research because it states that data access is based on assigning value to categories of data and analyzing both the risk and business need for each data category, providing definitions of acceptable use and levels of privacy/availability. In order to do this, an organization must learn what data exists, define the criticality of the data, and gain an understanding of how the data is used. With regard to implementation, the authors state there is a continuum of decision rights that range from centralized to decentralized that needs to be determined for each organization, but suggest a committee be employed to clarify the role of data and aid in these decisions.

Kooper, M. N., Maes, R., & Lindgreen, E. E. O. R**.** (2011). On the governance of information: Introducing a new concept of governance to support the management of information. *International Journal of Information Management*, *31*(3), 195–200. http://doi.org/10.1016/j.ijinfomgt.2010.05.009

**Abstract.** Information governance as an approach to better govern the use of information within and outside an organization is rapidly gaining popularity. A common and scientific ground for this approach has not yet been formulated. In this article the authors

describe a definition for information governance, extending the common, one-dimensional approach into a more generic statement. Starting from the well-known principles of IT governance the authors further explore the aspects of both information and governance. Four hypotheses are proposed to give ground to the use of information governance. These hypotheses will be the basis for further research.

**Summary.** This article contributes to this research primarily due to the importance it gives to the accessibility of data, but also in the fact that it details how a proper governance process can help make data more accessible. Kooper et al. state that their "fundamental belief (and premise) is that organizations with an instituted information governance process are more effective at seeking, collecting, processing and applying information, and are getting more value from their and others' information sources" (p. 195). It is stated that a major limitation of common forms of IT governance is that they exclusively exists to enable "control" and ignores the vital need for innovation, business development and value creation. Kooper et al. state that the focus should be on the use and exchange of information. They argue that the value of information is based on its meaning and interpretation within its context and that "governance actors" influence the interaction between the creator and receiver of information and facilitate making sense of the information.

Kooper et al. discuss a co-governance approach that allows interactive parties to collaborate and govern without a central or dominating governing actor. The authors suggest that this sort of approach may lead to a higher willingness to exchange information and also improve its reliability due to an agreed set of rules rather than a dictated set.

Meyers, C. (2014). How data management and governance can enable successful self-service BI. *Business Intelligence Journal*, *19*(4), 23–27.

**Abstract.** How can you build a dashboard quickly and efficiently without the risk of using inaccurate or outdated data? This article examines the role of data management and data governance in ensuring accurate and trustworthy reporting in a business intelligence solution. We introduce the power user model and explain how implementing data controls can strengthen the business-IT partnership.

**Summary.** Without directly stating it, Meyers examines the need for managing data in a way that focuses on accessibility and security. She discusses how disconnected data, such as un-moderated Microsoft Access databases that live in various locations in an organization make reporting quick and easy compared to data housed in IT, but because they have no control mechanisms they are not necessarily trustworthy for making good use of the data. This is valuable for this research because it indicates that simply more accessibility, in the exclusion of security and some controlled management, is not the best way to solve business problems either.

Meyers states that with a model that leans too far toward self-service, an organization will have a variety of users with various skill levels and understandings working with data and potentially coming to inconsistent or inaccurate results that will in turn lead to a distrust of the data in general. She suggests that there are certain elements of data management that should be put in place in a centralized manner and that a group of power users within each business line or department is a feasible solution to help end users with self-service requests. This balance works because it is reasonable that these power users can be trained in the organization's standards and yet can serve as a bridge

between the centralized authority and the business needs. Meyers also suggests that maintaining reasonable reporting data sandboxes helps in the control of production data because they can be used for general needs without all the necessary controls of production databases.

Mont, J. (2015). Data governance 101: Getting started. *Compliance Week*, *12*(137), 50–60.

**Abstract.** The article focuses on the use of Data Governance 2.0 in private business enterprises. According to Jeffrey Ritter, technology consultant and lecturer at Georgetown University Law Center, the only data worth investing in is the information that creates velocity for the business. It adds that protocols related to the administration of data should be developed through the cooperative efforts of the management, legal and compliance people.

**Summary.** In order to ensure security of data and make it appropriately accessible, the organization first needs to know what data it has and ensure that the data is cataloged. This article adds value to this research because it suggests that this effort cannot be offloaded onto IT because the business knows the data they need and the regulatory risks associated with the data, and that these efforts should be a cooperative effort. Mont points out that organizations can aid this effort by creating a strong chief data officer who can help navigate obstacles associated with getting multiple business areas on the same page, and the Garner statistics predicts that by 2017, 50 percent of companies will have done so.

Suer, M., & Nolan, R. (2015). Using COBIT 5 to deliver information and data governance. *COBIT Focus*, 1–6.

**Abstract.** The article discusses on the Control Objectives for Information and Related Technology (COBIT) 5, which provides guidance for business leaders and information technology (IT) practitioner for IT governance and information and data management. It mentions the function of Chief information officers (CIOs) in securing the data and information systems, data integration, and the established information life cycle of COBIT.

**Summary.** Suer and Nolan state that COBIT 5 provides guidance for IT practitioners and business leaders regarding governance and management of data and information, as well as a set of recommendations that allow the transformation of data to into information and then knowledge that has enterprise value. There must be a close cooperative relationship between IT and the business where the business owns the data and management of it while IT owns the process and technology for ensuring data are secured and yet available to the business.

Beyond the obvious needs in terms of data security at a technical level in its digital storage locations and transmission, ensuring data can be accessible while maintaining security all starts with defining enterprise data dictionaries and creating data classification schemes to provide a common understanding of all the data. This includes detailed information about the data ownership and appropriate data security needs for each category area.

This article adds specific value to this research not only because it validates points made by other articles, but because it notes how this effort also requires a standardized data architecture. That means data are integrated and transformed in a consistent way and into integrated environments when possible that avoids excessive or repeated data

manipulation in order to meet the business needs. By addressing the data architecture and gaining a detailed understanding of the data elements, their business value and classifications, data can be made adequately accessible and security. It is recommended to take this on in manageable chunks, picking data that deliver the biggest business value first.

Swoyer, S. (2016). IT's evolving role in data governance. *Business Intelligence Journal*, *21*(1), 49–55.

**Abstract.** Enterprises need to balance the needs of governance-privacy, sensitivity, compliance, liability, auditability, and traceability-with the business users who consume, analyze, interpret, and act on that data. In large part, that has meant a change in the role of governance and in IT. This article discusses the evolution of governance and IT's role as it has moved from denying and excluding access to data to an active, responsive, and accommodating force interested in ensuring data is easily consumable and fit for purpose.

**Summary.** Swoyer states that business users are now demanding access to data rather than allowing it to sit locked up in an IT realm, but acknowledges that with that change there is an even greater need to find a way to manage it, govern it and control it. It is discussed how the past emphasized highly centralized, rigidly structured, top-down control over all the data because that was the only thing possible at the time.

In this regard, Swoyer strongly states that in terms of data itself, decentralization and using disparate, diverge datasets is bad for an organization. This makes data harder to control and increases risk. Despite this warning, Swoyer suggests that over control of data in an attempt to create the desired "single version of the truth" is bad because he believes that notion overall is a fallacy that can never be achieved, and in many cases

should not be achieved due to the different purposes people have that seem like they are creating different numbers based on different needs.

Swoyer argues for a new method of do-it-yourself (DIY) self-service. This model takes IT out of owning, managing and controlling access to data simply based on a fear of bad things happening, and favors a governance model that is of the people, by the people, and for the people (where the people means the business areas) and actually argues that he does not believe there has ever been an authentic attempt to balance the needs of privacy, sensitivity, compliance, and liability with the needs of the people who consume, analyze, interpret and use information, and that it is time to try.

Beyond talking about this balance, the main value of this article is in how Swoyer points out that the main thing that really needs to be controlled is personally identifiable information (pii), and that business needs seldom require that risky pii. He believes that by masking the sensitive data and making the other data more accessible many business users can satisfy their needs without the risks posed with pii.

Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: Toward a theory of information governance. *Journal of Management Information Systems, 30*(3), 141–178.

**Abstract.** In recent years, chief information officers have begun to report exponential increases in the amounts of raw data captured and retained across the organization. Managing extreme amounts of data can be complex and challenging at a time when information is increasingly viewed as a strategic resource. Since the dominant focus of the information technology (IT) governance literature has been on how firms govern physical IT artifacts (hardware, software, networks), the goal of this study is to extend the

theory of IT governance by uncovering the structures and practices used to govern information artifacts. Through detailed interviews with 37 executives in 30 organizations across 17 industries, we discover a range of structural, procedural, and relational practices used to govern information within a nomological net that includes the antecedents of these practices and their effects on firm performance. While some antecedents enable the speedy adoption of information governance, others can delay or limit the adoption of information governance practices. Once adopted, however, information governance can help to boost firm performance. By incorporating these results into an extended theory of IT governance, we note how information governance practices can unlock value from the ever-expanding mountains of data currently held within organizations.

**Summary.** Tallon, Ramirez and Short state that information governance has two goals that include: (a) maximizing the value of information to the organization by ensuring it is reliable, secure and accessible; and (b) protecting information from human error, loss of timely access, inappropriate use or misadventure. This article is relevant in that it reiterates the point that creating proper data architectures will help avoid business areas creating silos, data duplication, and a lack of integration that may seem like it increases accessibility, but in fact grow data in a negative way and produces more opportunity for security issues.

This does not mean that all data needs to be fully integrated and centrally controlled. Tallon et al. state that allowing researchers to self-administer their data files and develop certain policies themselves can still allow for a secure and accessible environment if there are some institutional structures set for them to work within. One

other big point of relevance to this research is that the authors bring up the value in having a high-level data map that shows all the major data areas and how they relate, as well the value of high-level policy documents that detail certain constraints that need to be adhered to and the roles for the individual data custodians, steward and owners.

Watson, H. J., & McGivern, M. (2016). Getting started with business-driven data governance. *Business Intelligence Journal*, *21*(1), 4–7.

**Abstract.** Watson and McGivern explain why data governance is not just the right thing to do but a cost of doing business. They offer several ways your enterprise can get started with this discipline. As a side benefit, adopting such governance will align IT, BI, and business closer.

**Summary.** Watson and McGivern detail the importance of providing high-quality, well-modeled, secure and accessible data to a business for BI analysis. They attempt to provide a starting point to implement data governance to help in this regard, defining data governance as consisting of "people, processes, and technologies used to manage, protect and use data so that organizations can leverage it as an organization asset" (p. 5). Their discussion on the balance and distribution of governance is relevant to this research because it is critical to realize that any one-sided approach is not going to lead to the correct balance between security and accessibility. The authors state that data belong to the business functions that create it, and that while IT stores and manages the operational elements of distributing access, the responsibility for data quality and therefore governance lies with the business units and therefore it is best-practice that data governance reports to the CEO or CFO rather than the CIO.

Top-down and bottom-up approaches to implementation are examined, and it is suggested that an organization would not simply use one or the other, but calibrate their approach to meet the business need. Management articulates business need and creates the charter in the top-down approach, whereas the people doing the work that understand the data focus on the tools and processes in the bottom-up approach. The key is to find the desired balance of the two for each individual organizational culture - realizing the goal is to improve data for business value by providing more transparency and accessibility while simultaneously ensuring protection.

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *Journal of Data and Information Quality, 1*(1), 1–27. http://doi.org/10.1145/1515693.1515696

**Abstract.** Enterprises need Data Quality Management (DQM) to respond to strategic and operational challenges demanding high-quality corporate data. Hitherto, companies have mostly assigned accountabilities for DQM to Information Technology (IT) departments. They have thereby neglected the organizational issues critical to successful DQM. With data governance, however, companies may implement corporate-wide accountabilities for DQM that encompass professionals from business and IT departments. This research aims at starting a scientific discussion on data governance by transferring concepts from IT governance and organizational theory to the previously largely ignored field of data governance. The article presents the first results of a community action research project on data governance comprising six international companies from various industries. It outlines a data governance model that consists of three components (data quality roles, decision areas, and responsibilities), which together form a responsibility assignment

matrix. The data governance model documents data quality roles and their type of interaction with DQM activities. In addition, the article describes a data governance contingency model and demonstrates the influence of performance strategy, diversification breadth, organization structure, competitive strategy, degree of process harmonization, degree of market regulation, and decision-making style on data governance. Based on these findings, companies can structure their specific data governance model.

**Summary.** While this article focuses on DQM, the relevance to this research is partly in how it points out the need for each organization to find the proper structure of control and reiterates the points that roles and responsibilities are important in this process, and that a cooperation is needed between IT and the business to get the balance right. The authors emphasize what the authors call a contingency model to data governance that attempts to allow flexibility in the way each organization needs to work. Weber et al. state that this model represents  two design parameters: (a) organization placement, and (b) coordination of decision making.

**Data Security**

Sen, R., & Borle, S. (2015). Estimating the contextual risk of data breach: An empirical approach. *Journal Of Management Information Systems*, *32*(2), 314–341. http://doi.org/10.1080/07421222.2015.1063315

**Abstract.** Data breach incidents are on the rise, and have resulted in severe financial and legal implications for the affected organizations. We apply the opportunity theory of crime, the institutional anomie theory, and institutional theory to identify factors that could increase or decrease the contextual risk of data breach. We investigate the risk of

data breach in the context of an organization's physical location, its primary industry, and the type of data breach that it may have suffered in the past. Given the location of an organization, the study finds support for application of the opportunity theory of crime and the institutional anomie theory in estimating the risk of data breach incidents within a state. In the context of the primary industry in which an organization operates, we find support for the institutional theory and the opportunity theory of crime in estimating risk of data breach incidents within an industry. Interestingly though, support for the opportunity theory of crime is partial. We find that investment in information technology (IT) security corresponds to a higher risk of data breach incidents within both a state and an industry, a result contrary to the one predicted by the opportunity theory of crime. A possible explanation for the contradiction is that investments in IT security are not being spent on the right kind of data security controls, a fact supported by evidence from the industry. The work has theoretical and practical implications. Theories from criminology are used to identify the risk factors of data breach incidents and the magnitude of their impact on the risk of data breach. Insights from the study can help IT security practitioners to assess the risk environment of their firm (in terms of data breaches) based on the firm's location, its industry sector, and the kind of breaches that the firm may typically be prone to.

**Summary.** Sen and Borle define a data breach as an "incident that involves unauthorized access to sensitive, protected or confidential data resulting in the compromise or potential compromise of confidentiality, integrity, and availability of affected data" (p. 315). They examine several common theories in order to seek correlations between things like the numbers of vulnerabilities in data security, stronger laws requiring disclosure of data

breaches, money spent on IT, and the length of time between data breaches. The types of data breaches specifically examined that most impact this research include: (a) sensitive information posted publicly, mishandled, or sent to the wrong party; (b) intentional breaches by an insider with legitimate access; (c) lost, discarded or stolen paper records; (d) lost, discarded or stolen mobile devices; and (e) lost, discarded or stolen stationary electronic devices.

Sen and Borle find that while there is a correlation between the overall number of vulnerabilities and the time between breaches, it was odd to realize that reducing the overall number vulnerabilities only appears to help a very small amount in reducing time between breaches in education. They also find an unexpected result when they determine that all industries actually see an increased risk of breaches when IT expenditures rise, suggesting that money is probably spent in the wrong areas – usually going toward hardware and software controls rather than including administrative and other physical controls. Stricter laws relating to disclosure of breaches have a significant impact on the education sector, indicating that these laws tend to change the way education institutions act to minimize breaches.

Understanding these correlations and the realities of security that are not obvious is certainly relevant when seeking to balance data security with accessibility within a higher education institution. Findings seem to show that while expanding legitimate access to data may increase the overall number of vulnerabilities, this only very slightly increases risk of breach, and spending money on administrative controls that would involve policy, procedural guidelines and training may be beneficial over simply

expanding technical controls of data. Also, seeking to align these administrative controls with laws such as FERPA can help minimize breaches.

**Data Stewardship**

Small, M. (2013). From data breach to information stewardship. *Network Security*, *2013*(10), 5-8. http://doi.org/10.1016/S1353-4858(13)70112-0

**Abstract.** Loss and theft of information from organizations continues to be a significant problem. Given the amount of attention to this issue and the wealth of standards and technology available, why do these leaks still occur and what can be done to improve matters? People would not treat money with the same disregard that they treat information and data. Taking care to look after property that is not your own is called stewardship, and what is needed is better information stewardship, explains Mike Small, a member of the London Chapter of the ISACA Security Advisory Group.

**Summary.** Small discusses that data breaches occur in a range of ways that come from both inside and outside an organization. He states that these breaches can be from attack or theft, but most often occur due to simple misuse and accident – and that the most likely area for security risk includes data on portable devices. Small quotes a Verizon report that stated how end user devices were the one thing most likely to be compromised during 2012.

In order to combat this without taking an overly controlling approach, Small looks at the COBIT-5 framework. He suggests that the term “information stewardship” is all about managing the data in a secure and yet accessible way, and suggests that human behavior is key to creating good information stewardship and that creating a culture of security is how to address the issue. Creating this culture involves: (a) changing the

perception of security; (b) creating information stewardship champions; (c) educating, teaching and mentoring; and (d) rewards and sanctions. This essentially means that security should be marketed in a positive way, that respected people within the organization are needed to champion information stewardship, the value of information is clearly communicated and supported by clear security guidance and training, and that management will visibly support security by rewarding those that abide by the guidance and sanction those that do not.

The value to this article is the understanding that building a culture is very important in order to ensure security can be maintained while making data more accessible. The key is to include everybody that touches data in the information stewardship. Involving a range of stakeholders in this way creates a culture where people in the organization understand the sensitivity of information and the ways information can be put at risk. Small suggests examining COBIT5 ISO/IEC 27001:2005 and ISO/IEC27002:2005 because these objectives are aimed to ensure the confidentiality, integrity and availability of information.

## Conclusion

The references in the Annotated Biography section of this study identify tools and techniques that help leadership within higher education institutions balance data security with the high levels of accessibility necessary to provide value to the organization and answer key business questions in a timely manner. Data security and accessibility are not the only two critical aspects addressed by data governance (Chapple, 2013; Petersen, 2012; Swoyer, 2016; Watson & McGivern, 2016), but are arguably the yin and yang of data governance.

Research indicates that specific data governance tools and techniques can indeed be utilized to help balance the previously outlined security and accessibility concerns. The key aspects that help with this purpose include: (a) gaining a true and clear understanding of the organization's data; (b) defining clear roles and responsibilities in terms of ownership, stewardship and use of data within the business areas; (c) appropriately de-identifying reporting data for business users; and (d) determining and instituting an adequate balance of centralized and decentralized data policies (Chapple, 2013; Hickson & Dowdy, 2014; Petersen, 2012; Swoyer, 2016).

It is quite clear that in order to ensure an organization's data are both secure and accessible, and organization needs to develop a full understanding of their data (Chapple, 2013; Hickson & Dowdy, 2014). This includes not only defining the data (at least determining and documenting what data exists and where it lives), but also classifying it into categories (Chapple, 2013; Hickson & Dowdy, 2014; Petersen, 2012; Suer & Nolan, 2015). Defining data gives an institution a common language that Chapple (2013) states is necessary in order to effectively share information and therefore make it accessible not only in terms of security, but in terms of use based on the understanding of what the data actually means. After defining the data,

categorizing it into classifications can not only be used to show its business use, but according to Petersen (2012) should also be used to define levels of security (such as public, non-public and sensitive). Through this sort of categorization it becomes obvious which elements can be widely accessible and what needs more restricted access (Petersen, 2012; Suer & Nolan, 2015).

Defining clear roles and responsibilities in terms of ownership, stewardship and use of data is another key step to balancing security with accessibility (Chapple, 2013; Hickson & Dowdy, 2014; Kelly, 2015; Petersen, 2012; Small, 2013; Watson & McGivern, 2016). Watson and McGivern (2016) state that while IT is responsible for storing and technically managing access, the business units own the data, and thus Khatri and Brown (2010) state that business owners must have an important role in managing access. It is through these roles that Petersen (2012) suggests that the access and security can be appropriately determined and therefore granted on a need-to-know basis that still allows for meeting business needs.

The Responsible-Accountable-Consulted-Informed (RACI) matrix, as seen in Appendix B, is introduced as a tool to help an organization start to document these roles and ensure that data definitions are properly determined (Chapple, 2013; Hickson & Dowdy, 2014). This matrix defines who is ultimately responsible for (and therefore who owns) each particular data element, who is accountable for ensuring accuracy of the definition, who is consulted to fully define each element, and who is simply informed of the definition and any changes to it in the future.

Swoyer (2014) determines that personally identifiable information (pii) is the main data that needs to be controlled and that businesses seldom require that risky pii. Therefore de-identifying reporting data for business users is determined to help considerably when attempting to expand access without sacrificing security (Petersen, 2012; Swoyer, 2014).

Beyond the more operational and structural work discussed so far, related policies need to be developed that explicitly define access and authorization to ensure security is consistent, transparent, explicit and understood, rather than viewed as a barrier (Petersen, 2012). Determining and instituting appropriate data policies in cooperation with key data people in the organization is a key for maintaining the delicate balance between security and access. Tallon, Ramirez, and Short (2013) find that policy documents that detail certain constraints can put order and clarity around the roles and responsibilities for the data custodians, stewards and owners. They also state how policies can help with the creation of new data silos that are off the radar, and the creation or retention of unnecessary data, that both cause additional security risks.

Tallon et al. (2013) quote an unnamed Chief Information Security Office (CISO) of Intel Corporation as saying: "when you take a lock-down approach to information flow and you over-govern it, you destroy value, and if you under-govern it, you won't maximize the value for business" (p. 167). Kooper et al. (2011) and Tallon et al. (2013) assert that organizations that institute an adequate yet balanced information governance process are more effective at seeking, collecting, processing and applying information and are getting more value from their and others' information sources in a way that promotes better overall organizational performance. The techniques in this study should help institutions begin to find that equilibrium by focusing on the most essential and basic elements of data governance that lead to a balance between security and accessibility.

## References

Chapple, M. (2013). Speaking the same language: Building a data governance program for institutional impact. *EDUCAUSE Review*, *48*(6), 14–16.

Evaluating Information Sources. (2014). *Center for Public Issues Education*. Retrieved from: http://ce.uoregon.edu/aim/Capstone1Perm/evaluateinfo.pdf

Family Educational Rights and Privacy Act (FERPA). (n.d.). *U.S. Department of Education.* Retrieved from: http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html

Hickson, J., & Dowdy, J. G. (2014). Organizational structures for data governance at community colleges. *Education Advisory Board: Community College Forum*

Kelly, M. (2015, June 8). The chief data officer in higher education. *EDUCAUSE Review*. Retrieved from: http://er.educause.edu/articles/2015/6/the-chief-data-officer-in-higher-education

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. http://doi.org/10.1145/1629175.1629210

Kooper, M. N., Maes, R., & Lindgreen, E. E. O. R**.** (2011). On the governance of information: Introducing a new concept of governance to support the management of information. *International Journal of Information Management*, *31*(3), 195–200. http://doi.org/10.1016/j.ijinfomgt.2010.05.009

Meyers, C. (2014). How data management and governance can enable successful self-service BI. *Business Intelligence Journal*, *19*(4), 23–27.

Mont, J. (2015). Data governance 101: Getting started. *Compliance Week*, *12*(137), 50–60.

Petersen, R. (2012). Policy dimensions of analytics in higher education. *EDUCAUSE Review*, *47*(4), 44–46.

Ping-Ju Wu, S., Straub, D. W., & Liang, T.P. (2015). How information technology governance mechanisms and strategic alignment influence organizational performance: Insights from a matched survey of business and IT managers. *MIS Quarterly*, *39*(2), 497–A7.

Rosenbaum, S. **(**2010). Data governance and stewardship: Designing data stewardship entities and advancing data access. *Health Services Research*, *45*(5p2), 1442–1455.

Sen, R., & Borle, S. (2015). Estimating the contextual risk of data breach: An empirical approach. *Journal Of Management Information Systems*, *32*(2), 314–341. http://doi.org/10.1080/07421222.2015.1063315

Small, M. (2013). From data breach to information stewardship. *Network Security*, *2013*(10), 5–8. http://doi.org/10.1016/S1353-4858(13)70112-0

Suer, M., & Nolan, R. (2015). Using COBIT 5 to deliver information and data governance. *COBIT Focus*, 1–6.

Swoyer, S. (2016). IT's evolving role in data governance. *Business Intelligence Journal*, *21*(1), 49–55.

Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: Toward a theory of information governance. *Journal of Management Information Systems*, *30*(3), 141–178.

Thompson, N., Ravindran, R., & Salvatore, N. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*, *32*(3), 316–322. http://doi.org/10.1016/j.giq.2015.05.001

Watson, H. J., & McGivern, M. (2016). Getting started with business-driven data governance. *Business Intelligence Journal*, *21*(1), 4–7.

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *Journal of Data and Information Quality*, *1*(1), 1–27. http://doi.org/10.1145/1515693.1515696

# Appendix A

## Definition of Terms

The following definitions provide the foundation upon which this research builds:

**Data governance.** The people, processes, and technologies used to manage, protect, and use data so that organizations can leverage it as an organizational asset with the ultimate goal of creating value for the organization (Watson & McGivern, 2016, p. 5).

**Data accessibility**. The ability for the right people to have adequate security rights and tools to access data they need in an efficiently and usable format in order to turn data into actionable information and knowledge.

**Data stewards.** Individuals who maintain ownership and responsibility for designated data elements (Hickson & Dowdy, 2014).

**Data breach.** An incident involving unauthorized access to sensitive, protected, or confidential data resulting in the compromise or potential compromise of confidentiality, integrity, and availability of the affected data (Sen & Borle, 2015, p. 315).

## Appendix B

| Sprint | Term | Campus Data Steward | Registrar's Office | Institutional Research | Student Affairs | Finance | Human Resources | Provost |
|---|---|---|---|---|---|---|---|---|
| 1 | Academic Degree | A | R | C | – | C | – | I |
| 1 | Academic Term | A | R | C | – | C | – | I |
| 1 | Academic Year | A | R | C | – | C | – | I |
| 1 | Active Student | A | R | C | I | C | – | I |
| 1 | Campus Residence Status | A | C | C | R | C | – | – |
| 1 | Degree Seeking Status | A | R | C | – | C | – | I |
| 1 | Enrolled Student | A | R | C | I | C | – | I |
| 1 | Student Classification | A | R | C | – | C | – | I |
| 1 | Student Continuation Type | A | R | C | – | C | – | I |
| 1 | Student Level | A | R | C | I | C | – | C |
| 1 | Student Time Status | A | R | C | – | C | – | I |

| Roles | |
|---|---|
| Responsible | Role that owns the definition and leads the effort to accurately develop it |
| Accountable | Role that answers for the compensation and correctness of definitions across the University |
| Consulted | Roles that have active input into the development of the definition; two-way communication during the definition development |
| Informed | Roles that are kept informed on the development of definitions, mostly in a one-way fashion at major milestones |

*Figure 1. RACI Matrix (Chapple, ,2013)*