Analytics Governance: Towards a Definition and Framework

Emergent Research Forum

Atiya A. Avery
University of Illinois at Chicago
aavery3@uic.edu

Kyle Cheek
University of Illinois at Chicago
kdcheek@uic.edu

Abstract

Analytics is an emerging competency and an evolutionary step beyond traditional reporting and dashboarding, making it both an asset and liability to organizations. Organizations are currently implementing advanced analytic capabilities which require additional considerations not addressed by existing governance frameworks. A distinct analytics governance framework would help address issues surrounding the policies and processes of the analytics function. These issues include human capital development, integration of analytics into the broader enterprise, and legal and regulatory concerns. In this paper, we move towards defining an analytics governance framework by devising a set of guiding principles.

Keywords

Analytics, Corporate Governance, IT Governance, Data Governance, Information Governance.
Section 1: Analytics as an Asset and Liability

Analytics is an emerging competency and an evolutionary step beyond traditional reporting and dashboarding, making it both an asset and liability to organizations. Analytics facilitates the realization of business objectives through the reporting of data, the creation of prediction models, and the analysis and optimization of business processes to enhance organizational performance (Delen & Demirkan, 2012). The current business environment is hyper-competitive, with highly commoditized technologies leading to homogenized product and service offerings. The last frontier of differentiation is for organizations to maximize the returns on their business processes, and this requires sophisticated analytics (Davenport, 2006). Organizations, from UPS to public school systems, have utilized analytics to derive deeper insights in order to enhance performance (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011; Davenport & Dyche, 2013). Organizations who self-reported that their use of analytics differentiates them from competitors in their industry were twice as likely to be top performers (IBM Institute for Business Value & MIT Sloan Management Review, 2010).

Governance consists of an organization’s internal policies and processes that enable organizational alignment, human capital performance, and adherence to the legal and regulatory environment. In the current business environment companies are implementing advanced analytic capabilities that require additional considerations that are not addressed by existing governance models (Consentino, 2014). A distinct analytics governance framework would help address issues such as human capital development, integration of analytics into the broader enterprise, and legal and regulatory concerns.

An example of why analytics governance is needed occurred with the retailer Target in 2012. According to the New York Times, Target was at the forefront of marketing analytics, pairing data that it collected from its customers with other readily available data sources in order to better segment and target consumers. Marketers discovered—for example—that, during pregnancy, customers were more likely to form new shopping habits than at any other point in their lives. However, retailers typically sent out promotional materials after a child was born, which was too late to capture the customer. Target decided that it needed to reach these customers early on in their pregnancies. By utilizing analytics, Target developed a pregnancy prediction model that predicted the likelihood of pregnancy based on the “basket of items” that a woman had purchased. If this pregnancy predictor indicated a high probability of pregnancy, Target would send out promotional mailers for baby items to the customer. Target ran the model against the recent purchases of women in its database. Target then began to send out promotional mailings to women who were likely pregnant. In doing this, Target violated no existing internal or external laws, regulations, or policies. However, it appears that the age of the female (and the sensitivities around that) was never taken into consideration for the model. In one case, a father became outraged when his teenage daughter received promotional material in the mail insinuating that she was expecting; however, Target was correct: the man later confirmed with his daughter that she indeed was pregnant.

In essence, Target was aware of a teenage pregnancy before even those closest to the mother knew of it. Since this incident, Target has implemented controls in its marketing approach so that situations like this do not happen again (Duhigg, 2012). Existing governance frameworks such as data governance, information governance, and information technology governance would not have prevented this incident from occurring. Next, we discuss the benefits and shortcomings of existing governance frameworks for the analytics function.

Section 2: Existing Governance Frameworks

We define data governance as “the framework of decision rights and accountabilities to encourage desirable behavior in the use of data” (Wende, 2007, p.2). Wende’s definition appears to be in line with other definitions of data governance presented by Cheong and Chang (2007), Newman & Logan (2006), and Cohen (2006). Government regulatory and compliance requirements, such as HIPAA or the various privacy acts, often dictate an organization’s data governance goals with the primary focus on minimizing liability associated with noncompliance. However, a data governance framework is best implemented when data is viewed as an asset, not a liability; in this way organizational integration and business intelligence can be gained. Figure 1 denotes Oracle’s data governance framework, which views data as an
asset. The framework is a sequential process beginning with the development of policy, standards, and strategy and ending with data warehouses and business intelligence; (Oracle, 2011).

![An Example of a Data Governance Framework](source: Oracle)

This process is clear and intuitive. However, in the current environment, a data governance framework may not be able to keep up with the extraordinary increases in the types and amount of data that are being generated. The “Big Data” currently being produced is best governed from the perspective of a lifecycle. Therefore, a more comprehensive platform for managing data becomes necessary to address all phases of the Big Data lifecycle. The information governance process is structured differently from the data governance process and views data from the perspective of a lifecycle. Debra Logan a research VP at Gartner defines information governance as “the specification of decision rights and an accountability framework to encourage desirable behavior in the valuation, creation, storage, use, archival and deletion of information. It includes the processes, roles, standards and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals” (Logan, 2010).

This definition is aligned with the well-cited academic definition provided by Kooper et al. (2011) in which information governance is defined “as a ‘logical’ alternative, focusing on the seeking and finding, creation and use, and the exchange of information, and not solely on its production” (p.2). Much of the research on information governance is focused on the long-term storage and management of personally identifiable information. U.S.-based academic and practitioner research is sorely lacking in this area.

Information governance appears to have a limited application to analytics governance. However, its lifecycle approach is a useful perspective from which to view an ideal governance framework. An even more comprehensive governance framework than data and information governance is information technology governance (IT governance). The research on IT governance is not as sparse as the research streams for information and data governance, and progress has been made in this area with high-quality academic research such as Sambamurthy et al. (1999), Xue, Liang, and Boulton (2008) Watson, Fuller, Ariyachandra (2003), and practitioner research from the IT Governance Institute. IT governance is concerned with the performance of IT and its ability to manage risk. IT governance has a particular focus on an organization’s value creation efforts in regards to performance management and strategic objectives. (De Haes & Grembergen, 2004).

A much-needed attempt at standardization is COBIT 5 which coins itself as “the only business framework for the governance and management of enterprise IT.” (ISACA,2015). A cursory analysis of COBIT 5 and the existing governance frameworks reveals that they do little to address the emerging regulatory and human concerns for a case like Target. Furthermore, organizations are not fully implementing their existing IT governance frameworks and may not do so anytime soon. Despite this and in the meantime, emerging issues with analytics need to be addressed. Next, we discuss the unique needs of the analytics function.

**Section 3: The Need for Analytics Governance**

The analytics function in many organizations is rapidly growing, and its considerations need to be addressed. After a review of the three related governance frameworks, we see that there is a need for an analytics governance function to address a variety of concerns that are unique to analytics. We loosely define these needs as integration, legal/regulatory, and human capital.
Integration

Compared to traditional reporting and dashboarding, analytics causes a paradigm shift within organizations that will require new behaviors. People will need to collaborate more, new processes will need to be developed, and managers and executives will need to trust the decision support that analytics will provide them (Schrage, 2014; Satell, 2014). Currently, many organizations build analytic capabilities for each separate business function, thus creating silos. This leads to an absence of standards for analytic technology procurement as well as the deployment of technical artifacts from analytic processes. In other organizations, analytic outputs are not well integrated into the broader organization. This leads to unrealistic expectations about the capability of analytics, thereby creating an environment in which analytics is viewed as a novelty rather than an investment that can garner returns for the organization. In addition, those responsible for analytics within the organization may have the belief that the organization’s existing data and models must be perfect prior to use in a formal analytic framework. The organization may opt to spend all of its resources on addressing quality issues in lieu of gaining deeper analytical insights (Deloitte CIO Journal Editor, 2012).

Legal/Regulatory

The relationships between analytics and the legal and regulatory environment can often appear murky, with some organizations choosing to self-regulate and others taking advantage of the loopholes afforded by the lack of regulation. In the case of Target, it chose to self-regulate and corrected its marketing strategy towards women whom it believed were expecting (Duhigg, 2012). In contrast, vendors, such as the now defunct ReLead and activeconversion.com, use publicly and privately available datasets to provide companies with the email address and phone numbers of customers who anonymously visit websites (Tanner, 2013). Prior to the shutdown of ReLead, it was estimated that over 1,000 companies utilized this service and may have violated their internal privacy policies (Tanner, 2013).

Human Capital

The IT function and the analytics function are two distinct skillsets. Analytics is a business competency, not a technical competency, and business users are the end users of the analytics. To say that analytics is an IT competency is akin to saying that finance is an IT competency because spreadsheets run on computers. Analytics is a distinct function and should be recognized as such. Analytics is also an inherently human skill set: McKinsey Global Institute predicts that, by 2018, the United States will face a shortfall of over 1.5 million workers and managers who can analyze big data.

The increased use of Big Data is expected to transform healthcare, retail and manufacturing industries completely (Manyika et al., 2011), thereby requiring changing skill sets. Unable to support a fully staffed and trained analytic bench, some companies have responded by supplementing their existing analytics functions with crowdsourced labor, leading to unpredictable consequences. In 2009, Netflix held a contest to develop an algorithm that would predict what rating a viewer would give a video or movie after watching it. The data were “anonymized” and thus did not have any identifying personal information. A group of researchers discovered that the publicly available winning algorithm and anonymized dataset could be paired with another openly available data source and were capable of re-identifying customers (Narayanan & Shmatikov, 2008). In addition, Netflix realized that the winning algorithm, while mathematically superior, was not feasible in its production environment (Holiday, 2012).

The considerations needed for human capital, integration, and legal/regulatory guidelines are not addressed by the existing governance framework. A distinct analytics governance framework has the potential to solve the following concerns:

- Providing diverse frontline practitioners with training and best practices for integrating analytics into their current decision-making processes.
- Protecting businesses and financial institutions from increased liability caused by enhanced and insightful analytic capabilities.
- Providing a framework and best practices regarding the deployment, integration, and management of data mining models and software into new and existing business processes.

Next, we discuss guiding principles for an analytics governance framework.
Section 4: Guiding Principles for an Analytics Governance Framework

We propose guiding principles for the development of an analytics government framework. These guiding principles first address the behavioral issues that keep the data governance, information governance, and IT governance frameworks from being properly implemented. Second, we argue that an analytics governance framework must co-exist within a system of governance frameworks and present a framework that is conducive to this. Our guiding principles are based on extensive background research from corporate white papers, academic publications, and frontline practitioner blogs. Our guiding principles for the analytics governance is inspired by the work of the frontline practitioners. Two such examples are Unilytics (2015) and Kwantyx (2015) who have developed frameworks which address the policy and process needs of the analytics function. Figure 2 is a description of the components for strategically handling data once it is inside the organization.

**Accountability:** Creating precise definitions about who does what, when, why and how.

**Accessibility:** Ensure that analytics and reports are accurate, available, timely, and of direct value to stakeholders.

**Community:** Representation from all relevant business units with full participation by business owners and technology teams.

**Uniformity:** Definitions and policies set for all analytics reports, terminology, KPIs & calculations, reporting cycles, information dissemination, and routine review of governance for enhancements.

Figure 2
Part 1 of an Analytics Governance Framework
Strategy for Data Handling

These components are useful for addressing the behavioral issues that keep the data governance, information governance, and IT governance frameworks from being fully implemented. Once these components are addressed, the organization can then focus on each of the distinct governance frameworks noted in Figure 3. We argue that an analytics governance framework should be embedded in a system of governance frameworks that consists of data governance, IT governance, results sharing governance and decision-making. We briefly describe each of the governance frameworks as follows:

- Data governance. The organization’s data governance framework, including cleaning and archiving data.
- IT governance. An IT governance framework, such as COBIT 5, would be implemented here.
- Analytics governance. This governance process is driven by predefined key performance indicators. It defines “what the analytics is supposed to deliver.”
- Decision-making governance. This notion of governance is focused on officially holding people accountable for making (or choosing not to make) data-driven decisions. This is the human element of analytics.
- Results-sharing governance. This governance process focuses on the communication of results across the organization.
We acknowledge that many organizations may already be informally implementing parts or all of these guiding principles. We also acknowledge that governance is difficult to implement in practice because many organizations are not even proficient with basic data governance. Where decision-making governance and results sharing governance frameworks are a function of organizational culture, they are even harder to implement than the data governance, IT governance, and analytics governance frameworks. We hope to better understand the ideal analytics governance framework by utilizing studies and survey data in future research.
REFERENCES


