Whitepaper
Data Governance Roadmap for IT Executives
Valeh Nazemoff
The Challenge

IT Executives are challenged with issues around data, compliancy, regulation and making confident decisions on their business performance based on the information they receive. Data is the key asset of an organization and the lifeblood. How does an executive guarantee that the data they need will be reliable, accessible, accurate, complete, and secure? Executives receive vast amounts of data on reports, but do not know how to instantly translate them to their relevant needs.

Data is important because executives need to address strategic objectives to portray what needs to be accomplished, what performance measures will be used to measure progress against targets, what strategic initiatives have been identified to make actionable and operational, and what method will be used to articulate their strategy and communicate strategic objectives to both employees and internal and external stakeholders.

Do executives then have instant financial and non-financial information about the organization at their fingertips? Do they know how they are currently measuring against targeted goals? Of all the measures received, do they know why the organization has chosen those specific measures? Do they have a dynamic illustration that can simultaneously represent the historical performance, future outlook, and negotiated priorities across departments consistent with the business?

What are the metrics that are important to the business, setting the right targets, and demonstrating and reporting IT’s contribution to the business? IT Executives often ponder on how they can demonstrate the value to the business. Many IT organizations have a wide range of metrics that they measure with executives receiving various reports, but these reports are not presented in a summary fashion that is most relevant to the executive’s need to be able to communicate to their customers the service performance provided and how well they are meeting their Service Level Agreements (SLAs). Another driving need is being able to communicate back to the customers the spending of planned versus actual, value-added services provided, reporting achievements, and methods to drive improvement.

The key is to capture the right data and to present it in business-relevant terms. Businesses usually do not need reports detailing server-uptime or how long it took for help desk to respond to their call. They want reports that show how IT is contributing towards their objectives. The problem is how to communicate and translate something technical into a meaningful and useful language for all parties to understand and relate to.

Data presented in a dynamic data visualization format (Figure 1 shown below) that can be drilled up and down in order to discover relationship and conduct root-cause analysis would resolve the many concerns in addition to addressing the governance of data.
In this paper we will cover the Importance of Data Governance, the Key Drivers to Data Governance, and Acolyst’s Methodology and Roadmap for Implementing Data Governance.

The Data Governance Solution

Following recent trends, we see more and more business users defining the applications they need to conduct day to day activities, the demand for speed and instant access to data, movement to the cloud, and planning for the next big event, causes IT Executives to set Data Governance framework in place.

Data Governance is the overall management of the availability, usability, integrity, quality, consistency, and security of the data employed in an enterprise. It helps organizations meet compliancy with legislative laws, regulations, and mandates such as Sarbanes-Oxley (SOX) Act, HIPAA, and the HITECH Act.

Data Governance is a component of Enterprise Data Management, providing and enforcing enterprise-wide data standards, common vocabulary, reports, and the development and use of standardized data and processes which also helps organizations to improve regulatory compliance. Several approaches are used to organize Enterprise Data Management efforts such as Business Intelligence (BI), Data Security & Privacy, Master Data Management (MDM), and Data Quality Management (DQM).
Importance of Data Governance

There are real world issues that arise due to the fact that a Data Governance framework was not set in place. These are issues that can deal with life and death scenario. Looking at healthcare organizations, can a provider be able to view trend in test results for a new drug that has been launched in the marketplace? What about the ability to have instant access to information relating to the patient diagnosis and medical history?

An effective Data Governance framework can help organizations manage data more efficiently. It provides consistent definitions, and measures and tracks the quality of transactional and analytical data used across the organization. It enables organizations to more easily integrate, synchronize and consolidate data from different departments and to exchange data in a common format allowing for faster decision to occur.

Furthermore, it coordinates communication between businesses and IT and provides insight into the data across the business applications through shared terms and report format. Businesses are able to coordinate activities due to standardized processes and access to enterprise-wide data, causing improved business intelligence reporting.

Costs are often reduced as well with the implementation of an effective and efficient Data Governance framework. Savings are achieved by reducing the number of IT applications and systems and standardizing the ones that remain. Replicated data stores throughout the organization are eliminated, and data cleansing costs
are reduced as a result of data quality and data integration. Data quality also causes capacity to improve due to proficient reporting and analytical capabilities.

**Key Drivers to Data Governance**

As identified by David Waddington in an article published in Information Management Magazine (Sep/Oct 2010), titled *Data Governance, MDM and Data Quality: Information Difference Research Searches How Organizations Tackle Data*, we see that there are 9 (nine) core business drivers for IT Executives to implement Data Governance:

1. To Support BI/Data Warehousing Initiatives
2. To Support an MDM Initiative
3. To Facilitate the Migration of Legacy Data
4. To Meet Compliance and Legislative Requirements
5. To Reduce Corporate Risk
6. To Improve Corporate Flexibility and Business Agility
7. To Support Operational Software Upgrades (e.g. ERP, CRM, etc.)
8. To Reduce Costs
9. To Support Handling of Mergers and Acquisitions

When looking further into the key drivers, we can see that when implementing Data Governance to accomplish a specific goal, it automatically overlaps with other initiatives that can be accomplished within the organization. For example, if a healthcare organization needed to meet compliance and legislative requirements (4th Driver Listed), this would mean that to establish Data Governance, they would need to migrate data from legacy systems into new systems and formats (3rd Driver Listed).

**Acolyst’s Methodology and Roadmap for Implementing Data Governance**

There are 6 (six) components of Data Governance that would need to be considered and addressed for implementation.

The six components of a Data Governance framework are:

1. Organization
2. Policies, Principles & Standards
4. Investigation & Monitoring
5. Gap Analysis
6. Tools & Technology

---

To implement a Data Governance program, the organization needs to make sure that representative participation and commitment from both IT and lines of business exist, there is senior level executive sponsorship from both, and active consulting practices to drive and champion Data Governance implementation are conducted. Within IT, a dedicated active owner of data assets across the enterprise needs to be established. We call this person, Lead Data Officer (LDO).

The LDO is tasked with:

- Setting and authoring the direction of the Data Governance initiative, aligning business and IT goals, and managing organization data as a strategic asset.
- Driving business priorities and compliance with regulatory mandates.
- Defining roles and responsibilities for data owners.
- Creating data policies, procedures and standards for the organization as a whole.
- Directing how the data should be used, managed, and monitored across the organization.

**Policies, Principles, and Standards**

A policy must be developed for enforcing data standards and governance procedures that specifies who is responsible and accountable for various segments and aspects of the data, including its accuracy, accessibility, consistency, completeness, and updating.
For example, a healthcare organization with several applications might capture the date and time of a procedure in a specified field, where another system allows for the system user to enter the date and time within the notes section of the system. This results in inconsistency, risk of missing data, and data integration issues that would need to be assessed.

Setting policy demonstrates the importance and value of the data within the organization. Data is the most important asset in an organization and without standards and quality, the organization does not function effectively and productively.

---

*Figure 4 – Five Core Data Decision Domains*

There are 5 (five) core data decision domains that set the policies, principles and standards of data, as shown in Figure 4. Data Principles establishes the direction for all decisions and sets the organization’s standards for Data Quality, which is how data is interpreted (Metadata) and accessed (Data Access) by users. Data Lifecycle are the decisions that define the production, retention and retirement of data.

**Processes, Practices, and Architecture**

Processes must be established and formalized to guide principles for how policies, processes, and standards are created, collected, modified, implemented, and distributed across the organization. Without formalizing the process, IT constantly finds itself demonstrating their value add to the business lines. This is due to the issue that IT departments do not set formal Service Level Agreements (SLAs) with business units since they are within the same organization and other internal concerns that might be on the horizon. Setting formal processes and practices continues to identify and document how the organization manages its data. They define how the data is “to be” stored, archived, backed up, and protected. Practice and procedures are also instituted to ensure compliance and government regulations and audits are met.

Data Governance processes and practices helps organizations face challenges of enterprise level data integration concerns and includes enterprise standardization for data and systems. Figure 5 shows the importance of the Data Governance program structure to meet these challenges.
Figure 5: Data Governance Program Structure

The Enterprise Data Model Standard shown in Figure 5 is best known as the Data Architecture. Data Architecture addresses how the data is to be organized, integrated and includes enterprise data standards, data models, data flow diagrams, mapping spreadsheets, data definitions, and a metadata dictionary, in addition to security and privacy measures.

Data architecture is essential to be addressed at the beginning to understand what the organizational data requirements are. Commonly used data concepts are identified, which helps prepare the organization for effective and efficient data integration.

Data Integration

Data integration involves the process to cleanse, transform, merge and enrich data that is merged from multiple sources—whether from merger and acquisition, as part of a Master Data Management (MDM) initiative, or for a data warehousing project. A data warehouse system is a popular solution for extracting, transforming, and loading (ETL) data for business intelligence purposes.

In addition to the normal loading and consolidation of data, data integration addresses error handling, scheduling, process restart capabilities, data administration, gaps in data and audit. It ensures data is integrated in the timeframes required by the business and outlined in the SLA.

There are many business drivers for data integration. Some of them include the following:

- Integrating disparate data sources as a result of mergers and acquisitions.
- Master Data Management (MDM), particularly the need for a central source of consistent product or customer data.
- Integrating data silos of information across the organization, synchronizing data from different systems that use different formats, field names, data characteristics, and ensuring consistent data sources for application system.
- Migrating data from legacy systems into new systems and formats. For example, merging and migrating a patient case management system within a hospital with a patient legal matter management system within the same hospital into a new consolidated system.
- Reconciling inconsistent or replicated terminology into a single data dictionary based on standard
definitions, data types, and table layouts for each data element.

- Reporting data in standard formats and with standard interpretations.

Data cleansing process needs to occur and involves some of the following considerations:

- Identifying the differences in the data model schemes (data types, data element, length, value, etc.)
- Validating the rules based on the roles of the business users and their processes
- Recognizing the duplication of data, behaviors, and functionality

**Data Quality**

Problems with data quality are not only from incorrect data, but also the inconsistency of data. By creating and managing data models from the source system and creating enterprise standards, consistency and quality can be improved in the target systems, and allows for organizations to monitor and improve data quality. The utilization of a data profiling tool allows the data to be assessed, identifying cross system data overlap, and making sure the information is consistent. Data profiling also collects metrics and statistics that tracks the effectiveness of Data Governance across the enterprise and helps highlight errors that require attention. The metrics establishes measures for monitoring information, performance, and actions to continually improve enterprise data quality. As part of continuous process improvement, updated Data Governance performance metrics would need to be applied to monitor for data quality improvement and compliance with policies and standards.

![Figure 6: CA ERwin Data Profiler collects key metrics important to Data Governance](image)

**Data Models**

To create an effective enterprise data structure, data modelers and architects work with business users to understand their data usage and requirements, as well as technical staff to create common database implementation standards. An organization gains the most advantage from data models created using common
standards and processes that are consistent with the enterprise's view of the data. The data model can be thought of as the foundation for expressing a shared vision and a tool used for communication. It allows for a view to be created inspecting the similarities and differences between the data elements.

The enterprise logical data model (LDM) is a technical representation of the business' view of its data. A roadmap is needed to help navigate the pathway to data integration success and set the appropriate priorities for data-related projects. The roadmap can be thought of as the GPS for the location of the information.

With the implementation of enterprise data models, critical business initiatives and capabilities are aligned with integrated data in a manner that sustains, leverages, and aligns the corporation's goals and objectives. Basically, it is a business modeling methodology for capturing, displaying and connecting data integration planning information.

**Figure 7: CA Erwin Data Modeler helps provide an enterprise view of core data assets**

**Metadata**

Metadata is structured information and business rules about data. The structured information can include data lineage, business rules, business term definitions, ownership/stewardship, transformation rules, data mapping, source systems, structures of data, systems of record, data currency, and data access.

A data model and model repository provides a central location for a documented data dictionary using common business term definitions, data element definition, data element data types, entity layouts, data domains, along with a directory of applications. The metadata allows for an examination of duplicate definitions, dissimilarities
of definitions, and identifying consistent inconsistency. The metadata then becomes a shared information knowledge sharing where the definitions, data types, entity layouts, and domains are published.

When establishing a data model repository to contain the metadata, there are some considerations to account for across the organization:

- Different applications and systems have been built using various platforms and databases.
- The data contained across the different applications and systems might not be stored in a standardized method. There might be different meanings, different data types, and different naming conventions.
- Various definitions applied. For example, the term Patient can mean an In-Patient for one system versus an Out-Patient for another.
- Some information captured is in manual format.

**Investigation and Monitoring**

The investigation and monitoring component within Data Governance often times overlaps during the data quality stage when data is being centralized in the data warehouse and a data profiler is utilized. To investigate and monitor the data would involve a few steps:

1. Identify the data quality issues
2. Prioritize the issues based on urgency, importance, dependency, and critical success factors
3. Conduct root cause analysis to determine and identify the probable cause of the data issue
4. Formulating a corrective action plan
5. Decide on the next steps
6. Implement the fix
7. Monitor the results

![Figure 8: Investigation and Monitoring Life Cycle Component of Data Governance](image)

Within Data Governance, there are many data management silos that would need to be considered and analyzed as part of the process. In Figure 9 below, we see some of the management silos that most traditional enterprise level organizations are comprised of.
Gap Analysis

Information derived from the early components of Data Governance are inputs into the Gap Analysis Matrix and Documentation. The Gap Analysis focuses on mapping the organization’s governance policies and processes against industry standards and best practices. This allows the organization to have an understanding of where their organization is, what their target needs to be, and addresses plans to get there. The LDO will then have a plan to identify the tools and technologies needed compared to what is in-house today, recommendations for best approaches in creating a standardized model that accommodates the requirements and organizational strategic objectives and initiatives, requirements for a “to-be” data architecture and enterprise information model, business impact for implementing Data Governance and possible actions to take to mitigate risk, and methods of effective communication and training with the business lines and stakeholders.

Tools and Technology

In addition to a possible Data Governance implementation tools such as a data profiler, data modeling, and model repository when considering and evaluating tools and technology to be utilized during and post implementation, tools to consider would be a workflow data management application that would be able to alert, track, notify, escalate and approve Data Governance standards and policies during the improvement and modification stage for maturing Data Governance within the organization.

Acolyṣt’s Workflow Data Management Application for Data Governance allows for a standard process of verifying and modifying a data and metric definition, ability to provide and document business justification for the change, and submitting it for approval. The workflow process collates all requests in each cycle and redirects them to be approved by the LDO and executive data and metric owner. Once the data and metric receives all levels of approval, the change is then implemented and tracked. Implementing the change request process has enabled a stronger sense of trust and accountability across IT Executives.
Organizational Change Behavior

Executives are mostly concerned with compliancy, audits, and making sure their strategic initiatives are implemented efficiently and effectively. Data Governance is important in not only making sure that compliancy with legislative laws, regulations, and mandates are met but also there is consistency with definitions, results, accuracy, storage, quality, availability, integrity, and security of the data. It is important to communicate this message across the organization.

Communication throughout the organization is important when standardizing and adopting Data Governance processes and policies. Identifying the audience for who will be impacted is the first step and should be identified at the beginning of the project. Second is to create an awareness message on how they are impacted, what the changes mean to them, and how it benefits them and the overall organization as a whole. The third step is to address the audience’s role within the Data Governance framework, articulating who is responsible for what, allowing individuals to gain a sense of responsibility and ownership. When individuals are held accountable it motivates them to be part of the success and adoption to standardizing on Data Governance. A Data Governance Communication Plan helps to organize and define the different types of communication necessary that is needed to add value to the organization.
Summary

The Data Governance program is best when implemented using agile practices. This enables organizations to streamline the process and communication to solve the many data issues that exist: inconsistent definitions of data and metrics, lack of timely results of data, duplicate storage of data, different results for the same metric, among others. An iterative approach allows for more success since it focuses on a fraction of the data issue or business unit and keeps the scope small and manageable. A successful project can then be expanded out to more successful projects where individuals want to be engaged and part of the success and the vision of the executive leadership team. Each small project allows for an evaluation and lessons learned to occur and documented in order to revise and modify the process for that specific organization. The trick is to make sure the smaller implementation projects are conducted through repeated cycles and incrementally so the momentum of the overall Data Governance program is not lost but instead gains quicker and measureable results.

Bio

Valeh Nazemoff serves as Vice President of Acolyst where she is responsible for working with client executives in meeting their strategic initiatives, identifying new and emerging technologies, and implementing industry standards and best practices.

To find out more about Acolyst visit www.acolyst.com.
For more information about CA ERwin solutions, visit www.erwin.com.