Strategy & Planning: Data Governance & Data Quality

April 30, 2017

In the era of big data and data science, most commercial and nonprofit organizations realize the potential power of data in accelerating results. The challenge that most face, however, is the constantly shifting landscape of changing customer expectations and the associated need to ensure accurate and accessible data that informs decision making to serve customers with diverse needs.

A well-rounded data governance process can help ensure the consistent quality, integrity, and usability of data. Additionally, building an understanding of member/customer needs and interactions with the organization helps close the gap between expectations and the value the organization delivers. In this paper, ORI shares valuable insights we have learned from our work with associations and highlights several of the more critical keys to success.

INFORMATION & DATA MANAGEMENT

Information and data management (IDM) is the set of related disciplines that aims to manage the data asset fully, from conception to retirement. While data management and data integrity protocols are composed of the numerous topics and processes presented in Figure 1, ORI finds that focusing on data quality and data governance is often an ideal starting point. Efforts invested in strengthening quality are critical to ensuring decisions are made based on accurate, timely, and relevant data. The associated work of maintaining high-quality data through appropriate data governance magnifies the return on investment by embedding the importance of data as a strategic asset into the fabric of the organization's structures and processes.
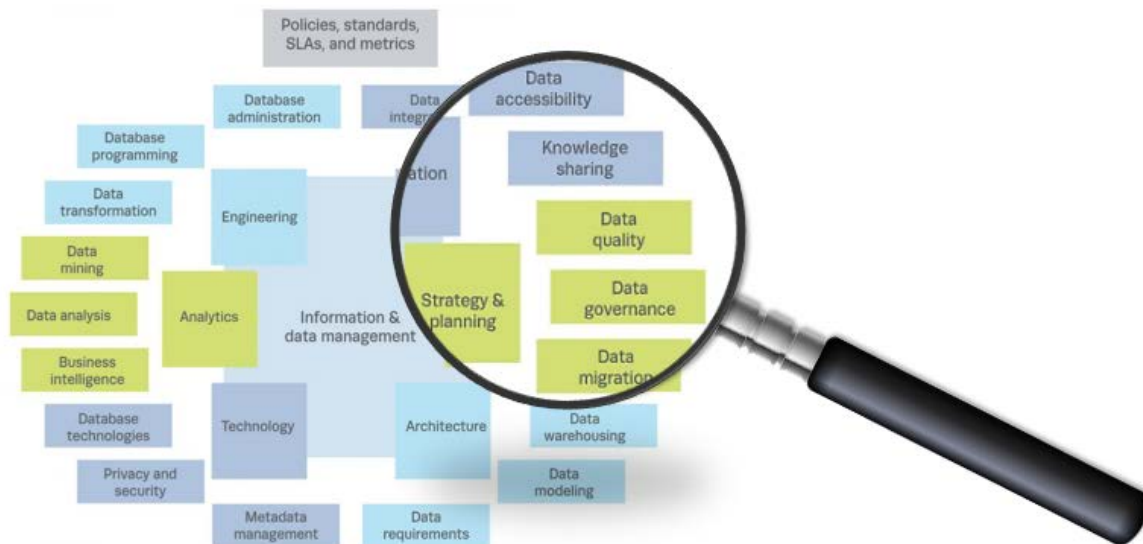

Figure 1: High-Level View of Data Management Disciplines (MITRE.ORG)

DATA QUALITY & ACCESSIBILITY

Determining which data truly matters to the organization is vital to any data quality initiative. The organization must identify priority fields early and narrow the focus to those deemed most important to critical business outcomes. Once priority fields are determined and agreed upon, a plan of action can be devised around the quality of this high-priority data. The related question then arises: What do we do with all of this other data? The organization simply cannot save everything and, as a result, analysts and data management personnel must fall back on judgment calls to determine which data to retain and how to store it—decisions that are much simpler when working with clearly identified priority data fields.

The following criteria and questions can be applied to determine which data to retain and for what reasons. Typically data should be retained if necessary for compliance or if any of the following questions are answered in the affirmative. Criteria may be weighted differently based on the long-term aims the data must meet as well as how highly it is valued in relation to those aims.

Actual Outcomes

A four-month data management and integrity project for a large industry association resulted in:

- Clear shared priorities for data improvements
- Healthier, more accurate databases
- Rigorous benchmarking processes to actively monitor database health, measure progress, and inform executive discussions and board presentations
- 50% gain in completeness of outreach data
- Robust data integrity protocols to collect, clean, integrate, and analyze data going forward
- Shared team perspective around data as an organizational asset and new ways to approach data management
- Increased confidence in existing data to support informed decision making

Quality Level
- **Description:** Is there enough information (e.g., from an up-to-date Data Management Plan) about what the data is, how and why it was collected, and how it has been processed to assess its quality and usefulness for the aims identified?
- **Quality:** Is the data quality adequate in terms of completeness, sample size, accuracy, validity, reliability, representativeness, or any other criteria relevant in the domain?

Demand
- **Known Users:** Are there users waiting for this data, or is there past evidence of a demand (e.g., will the data add value to an established resource or series)?
- **Recommendation:** Does the funder or a learned/professional society or equivalent body in the research field recommend sharing data of this type or on this research theme?
- **Integration Potential:** Does the data describe elements that fit standardized terms or vocabularies in other research domains (e.g., geographic locations and time periods)?
- **Reputation:** Was the data produced by a research group or project rated highly on the originality, significance, and rigor of previous research outputs? Will making the data available significantly enhance a group's or project's reputation?
- **Appeal:** Could the data have broad appeal (e.g., in relation to a landmark discovery, a significant new research process, or international policy and social concerns)?

Replication
- **Non-Replicable:** Would reproducing the data be difficult, costly, or impossible (e.g., as in the case of unrepeatable observations)?

Usage Barriers
- **Cleared:** Is the data classified according to its sensitivity and free from privacy/ethical, contractual, license, or copyright terms and conditions that restrict public access and reuse? Are any restrictions normal for the study domain?
- **Open Format:** Is the data in a format that does not require license fees or proprietary software or hardware for reuse?
- **Independent:** If any specialist software or hardware is needed to use the data, is it widely used in the field of study and readily available?

Storage
- **Unique:** Is this the only and most complete copy of the data?
- **At Risk:** Is the data held in a location that guarantees long-term storage?

DATA GOVERNANCE

ORI's research shows that member engagement strategy—along with the associated data requirements underlying customer experience improvement initiatives—is a top priority across the majority of senior association executives surveyed. In our conversations with clients, these priorities lead logically to the conclusion that data can no longer be viewed as the responsibility of one single department. Rather, data must become a responsibility that cuts across multiple groups. How one manages the data leads to several important decisions around data governance.

According to the Data Governance Institute (**DGI**), data governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models that describe the numerous stipulations of the what, who, when, why, and how of all data-related matters. Data health or quality is a key aspect of governance, and this decision-making system allows the organization to maintain its data health levels to ensure that critical strategy and planning goals can be realized. Data governance can entail:

- Organizational bodies
- Rules
- Decision rights (how we "decide how to decide")
- Accountabilities
- Monitoring, controls, and other enforcement methods

Data governance programs can differ significantly depending on their focus (e.g., compliance/security, data integration/migration, master data management). Regardless of the "flavor" of governance, however, every program will have essentially the same three-part mission:

- To make/collect/align rules
- To resolve issues
- To monitor/enforce compliance while providing ongoing support to data stakeholders

DATA GOVERNANCE AUTHORITIES

Somewhere in the organization, one or more groups is authorized to make rules and key decisions regarding data management. This group of individuals (or a hierarchy of groups) typically represents a cross-section of stakeholder groups. Together, these groups define a set of rules in the form of policies, standards, requirements, guidelines, and/or data definitions.

In ORI's experience, the most important aspect to consider when establishing data governance "authorities" is how best to engage key individuals across the organization in a regularly scheduled high-value discussion around data as a strategic asset. Connecting the users of the data (e.g., membership directors, marketing and communication team members) to those in IT and information management can reap huge benefits in terms of improving communication flow and understanding how the data is being utilized, what needs to happen to improve data quality, and how every team has a role to play in strengthening data quality.

DATA STAKEHOLDERS

A data stakeholder is an individual or group that could affect or be affected by the data under discussion. Most programs likely have a list of usual suspects—obvious stakeholders such as certain business groups, IT teams, and data architects/analysts.

However, other stakeholders may not be so obvious for a given decision or situation. Knowing which stakeholder to bring to the table—and when—is the responsibility of the Data Governance Team. Involving different individuals and connecting various end users aids in cultivating high-value discussions around data as a strategic asset for driving business objectives.

Data governance brings together cross-functional teams to make interdependent rules, resolve issues, or provide services to data stakeholders. These cross-functional teams likely come from across the organization but generally come from the business side of operations. They include groups that create data, those who use data, and those who set rules and requirements for data—setting policy that IT and data groups will follow as they establish their architectures, implement their own best practices, and address requirements. Data governance is the overall process of making this work.

One type of data stakeholder is a "steward." Stewards are concerned with taking care of data assets that do not belong to the stewards themselves. They represent the concerns of others, and some may represent the needs of the entire organization. Others may be tasked with representing a smaller constituency, such as a business unit/objective, department, or even a particular set of data. In most cases, stewards are senior representatives of stakeholder groups, and they convene to make decisions about the treatment of data assets.

ONGOING DATA GOVERNANCE PROCESSES

Every organization will decide how much structure, rigor, and formality to bring to the process of governing data. The Data Governance Institute (DGI) recommends formal, documented, repeatable procedures for:

1. Aligning Policies, Requirements, and Controls
2. Establishing Decision Rights
3. Establishing Accountability
4. Performing Stewardship
5. Managing Change
6. Defining Data
7. Resolving Issues
8. Specifying Data Quality Requirements
9. Building Governance Into Technology
10. Stakeholder Care
11. Communications
12. Measuring and Reporting Value

No matter how formal or rigorous the data governance process, it is critical to establish benchmarking procedures in order to actively measure and report progress in strengthening data quality as well as aid communication around data initiatives. In our work with clients to develop and implement strict protocols to

maintain ongoing data integrity, ORI draws upon a robust audit and benchmarking process to enable clients to actively monitor the health of their data, measure progress, and reinforce the concept of data as a critical organizational asset. Investments in improving data quality make it easier to identify trends and increase confidence in the inferences and conclusions organizations draw—ensuring that strategy is grounded in a reliable data foundation.

CONCLUSION

Strengthening data quality and ensuring the necessary governance processes are in place to maintain those data improvements over time is a key first step to understanding customers/members and what they need. Higher data quality also allows associations to track progress, improve real-time decision making, and justify expenditures to the board.

By connecting key strategic goals to the data required to make those goals reality, association executives can shift the dialogue about member engagement from a theoretical discussion with a high degree of unknowns to a precise conversation around business analytics, what needs to change, and how member needs can be met efficiently and effectively.

ORI is a research and data analytics firm that focuses on strengthening customer relationships. We offer quick and efficient data collection, agile research and analysis, and high-value team support and development. We transform data into decisions and decisions into stronger customer and member engagement. ***www.ORIresults.com***

For more information, contact ***info@ORIresults.com***