



DATA GOVERNANCE & DATA MANAGEMENT

Kate Nimety, Senior Director

ZURI GROUP

THE CASE FOR DATA STEWARDSHIP

Data is an asset to any organization. It represents the building blocks of our decision-making and can be modeled to help us understand where we are at present, as well as identify future risks and opportunities.

It also has associated cost which few of us measure. Duplication of effort, data entry inefficiencies, and loss of data integrity represent opportunity costs that all contribute to a loss of confidence.

And that loss of confidence affects more than just our ability to deliver consistent reporting internally. Our constituents trust us with their information - some private or sensitive - so we have a responsibility to be good stewards of the data that we collect and store.

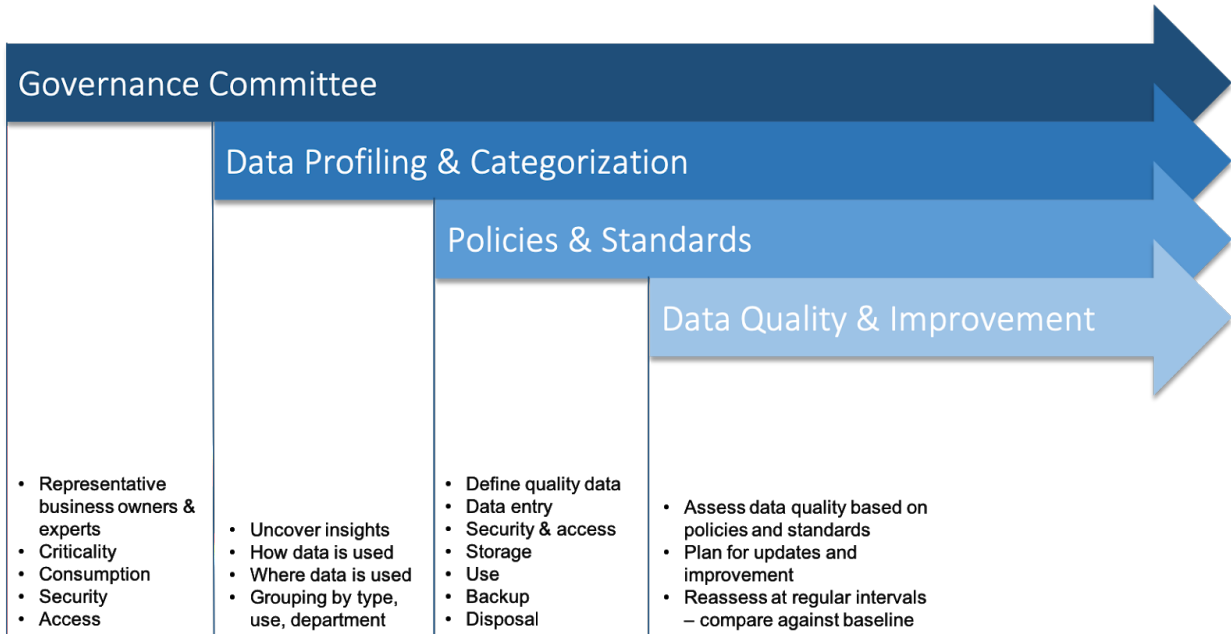
As toolsets become more sophisticated and our data needs mature, accuracy is critical. We need a framework that can allow us to establish a threshold for acceptable data quality to maximize ROI (Data Governance) and develop tools and processes that will help us maintain that quality, without overspending to attain perfection (Data Management).

DATA GOVERNANCE

Governance **defines who has jurisdiction** over our data, and **how we must capture and consume** that data *to ensure data quality*.

Successful Data Governance avoids isolated decision-making. Instead, it's active and collaborative, engaging a recruited committee of subject matter experts and representative data owners that:

- Profile the data
 - Identify critical data
 - Choose a methodology for capturing and maintaining data
 - Assess opportunities for data enhancement
- Categorize the data
 - Identify what other systems, tools, or reports access the data
- Define data quality or data health
- Outline policies and standards
 - Generate standards and policies that consider the security, accountability, and integrity of each set of data
 - Review requested changes to those policies which may have downstream impact. *(See Data Governance Committee Charter)*
 - Ensure consideration and understanding of impact before changes are made
- Drive communication to the user population
- Meet regularly to review standards and policies



Profile the data

The Governance Committee will perform Data Profiling to uncover insights which help define the access, use, and impacts for each category of data. Though not a full analysis, a data profile will help you understand how your data is behaving now, and even how business processes have been built around the data today. This becomes input for establishing policies that optimize data collection and maintenance.

Choose a methodology

Understanding data ownership, primary source systems, and data consumption across departments will help you determine which of these models is best for your organization.

Centralized: Master Data Management (MDM) is a method that centralizes critical data used across the entire organization and applies governance policies and standards- ensuring a single source of truth and consistency for the entire organization.

Decentralized/Distributed: Some organizations maintain data in a decentralized landscape of source systems, but recognize which system is the

source of truth and accuracy for each category of data and apply that methodology to data policies.

Categorize the data

The Governance Committee will use insights gained from data profiling to categorize and sub-categorize data based on type, relationship, department-specificity, use, security requirements, or source system. This will become your data taxonomy- the foundation you'll use for establishing policies.

Group information together in ways that help support the development of standards and policies. Biographical data, for example, is handled and consumed differently than revenue or financial data.



Define data quality

Understanding what makes information 'healthy' or 'high quality' sets the metric for acceptability. Defining what quality data means, categorically, for your organization will make it easier to see anomalies and outliers.

Outline policies and standards

A data policy defines the standards for data quality, as well as the full lifecycle for each category or sub-category of data: access and security, storage, protection, backup, and disposal.

The policy will also identify:

- How the information should be entered
- How the information should be used
- Which systems, outputs and reports and the various systems will use the information

This will help to detect impacts if something changes within the data or with the way it is structured.

Additional consideration must be made for information that is included in a centralized reporting tool or data warehouse, as the transformation logic applied may alter the way we consume the same data from one system to another. *(See Data Governance Policy Template)*

Drive communication

Sometimes, it will be necessary to communicate progress and changes to your user base. Whenever possible, engage the Governance Committee to identify how best to communicate status and timelines and incorporate policies and data management procedures into system training.

Meet regularly

As your organization's needs grow and change, your Governance Committee should meet to review existing policies and standards, making revisions and amendments as needed. Setting up a regular meeting cadence ensures ongoing attention and elevated priority of Data Management.

DATA MANAGEMENT

Data Management is the process of **analyzing data sets against governance policies, making updates to comply** with those policies, and **regularly monitoring the data sets for quality and health**. Once a governance framework is established, the work of data management can begin.

Oversight

In the **Centralized** methodology, Data Management falls under the oversight of a single unit or team. This team is responsible for establishing the tools and processes needed to collect and maintain healthy data.

In the **Distributed** methodology, Data Management is the responsibility of each department to collect and maintain the data they own. Oversight is required to facilitate cross-department communication, and to ensure that policies are upheld across the organization.

In either methodology, Data Management includes:

- Identifying a baseline for data health
- Performing regular data quality assessments
- Defining processes to collect and maintain accurate data
- Setting expectations for the threshold of acceptable data quality
- Enlisting the data owners to review assessments and advise policy updates
- Assessing impact of data structure changes

Identify a baseline

When initially launching a Data Management practice, it is important to establish a baseline of data health. This baseline will become a point of comparison as data quality improves. Analysis should begin by comparing current data against the expected level of accuracy or completeness as defined by governance policies.

Using the business rules set forth in the policy documents, review data categorically to understand:

- What records are considered VIP or Front-of-the-Line, and should always meet acceptable standards
- Percentage that is compliant with the policy – data health indicator
- Percentage that is outlier, but still acceptable – this may indicate a change is needed to the policy to accommodate these scenarios
- Percentage that is outlier, and must be updated:
 - **Update Now:** If the updates are easy to identify and apply or if data consistency will make an impact on fundraising in the near term
 - **Update During Conversion:** If the updates require combination or parsing of existing data elements, if historical data does not match current data, or if existing reporting would change after the data is updated
 - **Update After Conversion:** If the updates are easier to perform in the new environment and operations can continue despite the data inconsistencies, or if the updates rely on combined or parsed data elements

Perform regular assessments

The queries or reports should be reusable so that the same assessment can be performed at regular intervals. If the first assessment is considered the baseline for comparative metrics, performing the assessment again after your data quality initiative has begun will help understand the positive impact of your investment of time and attention. Setting up a Data Health Calendar to revisit the assessment with departments or by category will keep this effort current and limit obsolescence.

Define processes

Once the overall baseline health is identified, your Data Management Plan should:

- Define processes to update and maintain data that support the policies and business rules set out by Data Governance

- Refine data entry processes to capture and maintain information consistently
- Define who is responsible for enacting the processes
- Outline the cadence for health checks and accountability goals

Set expectations

Acknowledge that your data may not be as complete or as accurate as you assumed, and that the time and effort to update records is a necessary -and worthwhile- investment. Whenever possible, engage the Governance Committee to define acceptable data health thresholds, data quality goals, and how best to communicate status and timelines to your user-base.

Enlist data owners

With the understanding of areas that need remediation to meet data quality expectations, enlist the help of the data owners themselves. Departments that have administrative oversight of source systems become the primary owners of the data their systems maintain and must have the authority and responsibility to steward the information they collect. The Data Management team can work with data owners to identify the breadth and depth of updates that need to happen, as well as resources that can help accomplish the defined data quality standards.

At this point, keep in mind that a team will expend 80% of their effort to cleanse the last 20% of data, when striving for perfection. Instead, strive to meet the thresholds set by the Governance Committee, and focus on the records which are the most important to perfect.

Critical data elements such as VIP constituents or Front-of-the-Line (FOTL) records often require special handling or direct attention when assessing data health. These records must be as accurate as possible to ensure that the donor, alumnus, or auditing body receives the most complete and correct information possible. Prioritize these records and opt for bulk-updates where possible for the remainder.

Assess impact

Before making a change to data structures, data entry fields, hierarchies, or relationships, assess the impact to other reports or systems in the landscape. Impact assessment should begin with the Data Management team, but must be brought to the Data Governance Committee ultimately for review against defined policies. Ideally, those policies already include a list of upstream dependencies and downstream impacts, which make an impact analysis much more straightforward.

If necessary, take the opportunity to review and record the other systems, departments, reports, and websites that use or display this data. Include desktop applications such as Excel and Microsoft Word merge templates. Even if the change appears minor, without a full impact assessment, it is unclear if adding a value or inactivating a field will cause a report to count differently or an automated merge to fail.

VALUE PROPOSITIONS YOU CAN USE

Data Governance and Data Management are tools that your organization can invest in to ensure that data - one of your organization's most valuable assets - is protected for years to come.

- Establishing a plan for collecting, maintaining, using, and updating information will support more consistent data and higher levels of confidence.
- Recognizing the high value that data has in measuring your organizational success will help to support the initial cost of implementing Data Governance and Data Management practices.
- Applying standards and analytic rigor contribute to responsible stewardship of your data and the ability to rely on it for insights and decision-making.

FROM THE AUTHOR

If you're ready to get started or have questions about where to begin- **Zuri Group can help!**

I've had the privilege of working with a number of nonprofit and higher education organizations to establish a Data Governance groundwork and a consistent, sustainable Data Management practice. I'm confident that I can help you, too.

Kate Nimety
Senior Director, Strategic Services
kate@zurigroup.com
zurigroup.com

DATA GOVERNANCE COMMITTEE CHARTER

Vision Statement: *What drove you to establish this committee? Why did you choose the members you did, and who do they represent?*

Mission Statement: *What will this committee focus on? What does the future look like when policies and standards are established and followed?*

Goals: *What are the near-term and long-term goals for a governance methodology? These can include goals for data accuracy, data management, and also the confidence users have in the data.*

Success Measures and Committee Management: *Outline the metrics that define success. These should be measurable, feasible, and time-based. What is the meeting cadence and standard agenda? What are the requirements to establish a decision?*

Member Roles and Responsibilities

Member Name	Title/Department	Committee Role	Responsibilities

DATA GOVERNANCE POLICY TEMPLATE

Data Governance Policy | *Name of Policy*

Date Created: _____ Version No.: _____
Created By: _____

Data Category: i.e. Biographical

Data Type: i.e. Contact records

Background: *Describe the data entry, maintenance, security, storage or accuracy concerns with this data. This should include the purpose of this policy.*

Policy Scope: *What data is included under this rule? What data is excluded? Declare VIP records that are acceptable outliers so that they do not get updated inadvertently.*

Policy Rules:

Behavior	Business Rules	Impacts
Access/Security		
Usage		
Integrity/Accuracy		
Storage/Backup		

Role-Based Business Rules:

Role/Department	Business Rules	Impacts

Business Owner:

Resources/Reports Using this Data:

Resource or Report	Resource Owner	Data Use
		<i>Parameter, output, used to calculate data concept</i>

Data Definitions:

Field Name	Table Name	Business Definition