Alation

Data Governance Methodology

Overview

Modern data governance leaders understand the importance of high-quality data and are searching for the role that governance should play in delivering high-quality data, and how to implement that approach to data governance. That may seem strange given the volume of available material from 'independent experts,' but the truth is, legacy approaches are complex, resource-intensive, and feel imposed on the business rather than for the business. In a word, the experts are failing and smart leaders are looking for a better answer.

Alation offers a new and fresh approach to data governance with a focus on building a community of business experts who are invested and committed to improving their organization's effectiveness and data literacy.

Key Characteristics

The differentiating characteristics of Alation's Governance Methodology are:

- **Pragmatic, not theoretical** – The starting point for what to govern is the reports that are used monthly, quarterly, and annually to run the business. By definition, those reports contain the most important business metrics and are supplied from the most important data sources. It is natural then that the most impactful governance activity is to maintain the completeness, quality, and accuracy of these data sources for the benefit of the entire organization.

- **Business, not IT-centric** – Alation's governance approach begins with business content and requires business ownership to sustain a high level of quality that delivers value to the entire employee population. This is beyond participation in a program that is led by IT, but ownership of the process itself as supported by IT.

- **Incremental not 'big-bang' implementation** – Alation's governance approach focuses on repeating a successful pattern of implementation across business areas. This enables the rapid delivery of value to the business, and for decisions to be made about incremental investment based on proven success.

- **Community driven not committee controlled** – Alation's governance approach empowers employees to participate in a community where they are welcomed and rewarded for collaborating, sharing their knowledge, and opinions.

- **Guided not gated participation** – Alation's governance approach is rooted in the belief that people want to do a good job and will follow guidelines for participating and contributing when trained. It is therefore in the greatest interest of the organization to not use the role of steward to block participation or meter the flow of content but to instead act as mentor, coach, and guide to help people do things right.

- **Employee value, not steward & IT centric value** – Alation's governance approach focuses on delivering value to employees; not creating a mechanism to drive value to stewards or IT and enabling them to have complete power to control the flow of knowledge.

- **Data usage, not documentation deadends** – Alation's governance approach goes beyond simply documenting and creating glossaries. It integrates the governance policies, standards, terms, etc., with access to data, queries, and reporting. This embeds and makes governance an active and natural part of the work activities of analysts and users.

Alation's approach is also unique because it defines data governance as three distinct components that can be implemented independently or together. This provides a level of clarity about what, who, and how to implement that allows leaders to make clear investment decisions and understand what results can be expected.

1. **Business Governance** – Maintaining the content, quality, and understanding of core business assets and their supporting data.

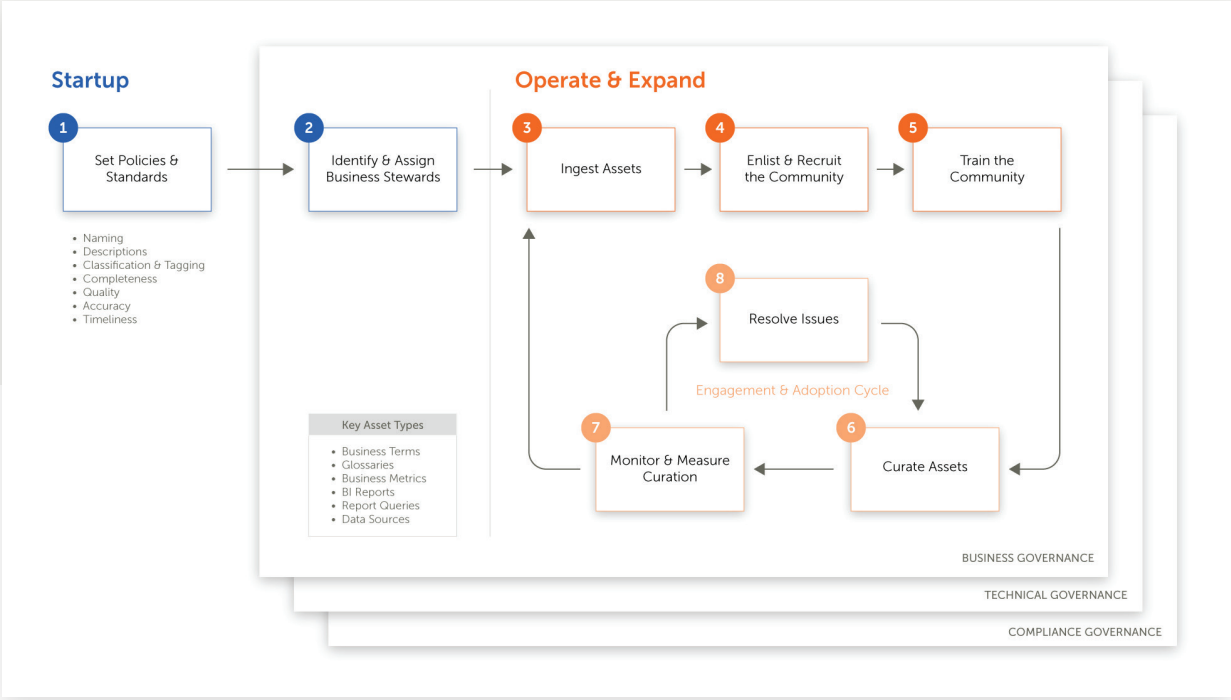
2. **Technical Governance** – Maintaining the content and quality of extended data attributes and data sources in the enterprise data landscape.

3. **Compliance Governance** – Maintaining the policies, classifications, assessments, and audit standards required for regulatory reporting.

Each of these is provided with a full complement of roles and responsibilities, asset types, policies, standards, and work processes. Table 1 in the Appendix depicts the required attributes for each.

Active Governance Process

Alation's approach to governance focuses on an iterative process of continuous management and improvement. The participants in the community are central to this process and both consume and layer in new knowledge while stewards guide them based on standards and policies.



1. **Set Policies and Standards** - The entire governance process requires that some things be decided in advance; namely, what will

be governed, how it will be governed, and how success or failure will be measured. These represent target commitments that are expressed as policies and standards that form a framework for the entire organization. They define completeness, quality, accuracy, timeliness, usage, access and classifications for both metadata and data. Governance programs often suffer from a reliance on dusty policy documents that are not easily accessed to somehow act as a part of everyday business activities. In contrast, Alation provides policies as fully searchable wiki-like pages, glossaries, taxonomies, and attachments to all associated data assets, enabling anyone to understand the rules when they are using the data. Users have immediate access to these as a guide as they search, use data, and participate by contributing their own knowledge to the catalog.

2 **Identify and Assign Business Stewards** - As implementation of governance begins for each business area, it's critical that a full-time business steward is identified. Generally, stewards sit inside the business area and play a critical role in driving awareness, onboarding, moderating and guiding participation. They are accountable for ensuring the targets in the policies and standards are maintained. Historically, governance teams have tasked stewards with primary responsibility for this enforcement work – they are the so-called policy enforcement officials of data. However, in the modern era, there's just too much data and far too much change for stewards to keep up. To address this problem, stewards and governance teams need to foster and build a community of willing participants who are recognized and rewarded for helping make data governance a key part of the organization's culture.

3 **Ingest Assets** - Governance requires knowledge and content. Historically, this knowledge has been gathered by assembling experts (stewards and SMEs) to identify key business assets on a one time basis. The complexity and ongoing change of a modern enterprise makes this approach unsustainable. Alation's approach is to have Business Steward's identify key assets (reports, terms, metrics, queries, and data sources) to be automatically ingested and analyzed. The analysis includes gathering behavioral insight, such as the most common users and popularity of an asset. Alation also uses machine learning to provide understandable names for abbreviated

and critically named data assets. Most critically, Alation continuously monitors these assets as the underlying people, process and systems change. As a result, Business Stewards can consistently and easily measure the gap between the governance standards and existing content using curation dashboards and analytics reports. New assets are discovered during automated updates, perpetuating an active cycle of assignment and continuous curation.

4 **Enlist and Recruit the Community** - This step is focused on building an ecosystem of participants and stakeholders who will adopt responsibility for contributing and maintaining the catalog. The keys to recruiting and building a strong community are: the clear identification of stakeholders, recognition and rewards. Alation identifies top users of specific data assets during the ingestion step, which aids the Business Stewards in the assignment of curation responsibilities. Recognition is based on usage and participation analysis and is used in communications that identify individuals and teams in order to celebrate their contributions. Recognition is also provided through more formal mechanisms, such as certification levels, badging, etc. There are two dimensions to rewards. The first is the personal satisfaction that each individual gets from experiencing how effective and useful it is having a trustworthy source of knowledge. The second is more tangible and can be provided in the form of bonuses (monetary, gifts, parties, time-off, etc.) associated with participation and maintaining quality standards.

5 **Train the Community** - This step is the beginning of a self-perpetuating cycle of engagement and adoption. The community is the key to driving a data culture that scales while following quality norms and enforcing behaviors and policies. New community members are onboarded, trained, and participate by sharing their knowledge and curating assets based on standards. This increases the breadth, depth and quality of information, which in turn attracts more new community members.

New community members learn the role of participants, usage policies and guidelines, catalog navigation, how to contribute, how to collaborate and governance assets standards and policies.

They also learn how they are recognized for their knowledge and contributions while also benefiting from the contributions of others.

6. **Curate Assets** - Curation is the process of populating, checking, and refining the attributes of an asset to provide users with greater trust through insight and knowledge. It is one objective of the governance process to maintain the required attributes for each asset as stipulated in the policies and standards. The Business Steward is accountable for maintaining the standards but is expected to do so through delegated responsibility to top users and other volunteers. The keys are training, recognition, and ensuring that participants realize that they have an ownership stake in the quality of the content.

7. **Monitor & Measure Curation** - Business Stewards have a responsibility to mentor and guide the community to build very high quality content. To do so effectively, they leverage curation activity reports, change notifications, and conversations to respond to questions and request assistance. Business Stewards also measure curation progress against the policies and standards using analytics and stewardship dashboard reporting. These capabilities allow the Business Steward to zero in on information that is missing and cross-check updates.

8. **Resolve Issues** - This step involves the Business Stewards requesting a change, receiving a change request, or an escalation. The key mechanism used to address all three is conversations. Conversations can be created on any topic and assigned to an individual and/or include a group of participants. As a result, Business Stewards can create specific requests for users as conversations, users can request things from the Business Steward, and groups can have moderated discussions to resolve conflicts. It is an incredibly flexible communication mechanism that allows for accountability to be tracked and enforced.

Appendix

Table 1 - Asset Attributes by Governance Dimensions

Assets	Governance Dimension		
	Business Governance	Technical Governance	Compliance
BI Servers	• Business Name (Title) • Business Description	• Technical Name • Source Type • URI Location • Owner Name • Technical Steward	• Accesses Sensitive Data • Accesses Restricted Data Sources • Compliance Steward • Related Policies
BI Projects	• Business Name (Title) • Business Description	• Technical Name • Project Type • Associated Report Source/ Server • Architecture Zone Classification • Access Classification • URI Location • Owner Name • Technical Steward	• Accesses Sensitive Data • Accesses Restricted Data Sources • Contains Access Restricted Reports • Compliance Classification • Compliance Steward • Related Policies
BI Reports	• Business Name (title) • Business Description • Usage Frequency • Report Type • Primary Report Users • Applicable Business Unit(s) • Key Report • Associated Briefing Book(s) • Associated Report Project • URI Location • Field List • Report Owner Name • Report Health Classification • Report Health Description • Certification Status • Date of Last Certification • Business Steward	• Technical Name • Associated Report Project • Architecture Zone Classification • Access Classification • URI Location • Owner Name • Technical Steward	• Accesses Sensitive Data • Consumes Restricted Data Fields • Requires Risk Assessment • Related Risk Assessments • Compliance Classification • Related Policies • Compliance Steward • Last Compliance Review Date • Compliance Review Notes
Data Sources	• Associated Metric(s) • Associated Query(s) • Business Steward	• Technical Name • Business Name (Title) • Business Description • Source Type • Architecture Zone Classification • URI Location • Owner Name • Technical Steward	• Contains Sensitive Data • Contains Restricted Data Sources • Supports Restricted Reports • Compliance Steward • Related Policies
Schemas	• Business Name (Title) • Business Description • Associated Metric(s) • Associated Query(s) • Business Steward	• Technical Name • Associated Data Source • Schema Classification • Update Frequency • Stewardship Classification • Access Classification • Data Origination Points • Primary Feed Mechanisms • Associated Feed Jobs/Assets • Associated Queries • Schema Quality Score • Technical Steward • Schema Health Classification • Schema Health Description	• Contains Sensitive Data • Contains Restricted Data • Compliance Classification • Related Policies • Compliance Steward

Assets	Governance Dimension		
Tables	• Business Name (Title) • Business Description • Associated Metric(s) • Associated Query(s) • Business Steward	• Technical Name • Table Classification • Update Frequency • Stewardship Classification • Data Origination Point • Primary Feed Mechanism • Associated Feed Jobs/Assets • Associated Queries • Table Quality Score • Technical Steward • Table Health Classification • Table Health Description	• Contains Sensitive Data • Contains Restricted Data • Compliance Classification • Related Policies • Compliance Steward
Columns	• Business Name (Title) • Business Description • Associated Metric(s) • Associated Query(s) • Business Steward	• Technical Name • Field Data Type • Most Frequent Values • Most Frequent Patterns • Range of Values • Uniqueness • Field Quality Score • Field Quality Rules • Field Quality Thresholds • Technical Steward • Field Health Classification • Field Health Description	• Sensitive Indicator • Sensitive Classification • Associated Regulations • Required Risk Assessment • Related Risk Assessments • Compliance Classification • Related Policies • Compliance Steward • Last Compliance Review Date • Compliance Review Notes
Queries	• Business Name (title) • Business Description • Query Statement • Query Author • Associated Report(s) • Associated Data Source Field(s) • Endorsements • Published Status • Steward	• Name • Description • Query Statement • Query Author • Associated Report(s) • Stewardship Classification • Associated Data Source • Endorsements • Steward	• Uses Sensitive Data • Uses Restricted Data Sources • Supports Restricted Reports • Related Policies • Compliance Steward
File Sources	• Associated Metric(s) • Associated Report(s) • Business Steward	• Technical Name • Business Name (Title) • Business Description • Source Type • Architecture Zone Classification • File Source • URI Location • Owner Name • Technical Steward • Source Health Classification • Source Health Description	• Contains Sensitive Data • Contains Restricted Data Sources • Supports Restricted Reports • Compliance Steward • Related Policies
Files	• Business Name (Title) • Business Description • Associated Metric(s) • Associated Report(s) • Business Steward	• Technical Name • Associated File Source • File Classification • Update Frequency • Stewardship Classification • Access Classification • Composite File • Data Origination Points • File Source • Primary Feed Mechanisms • Associated Feed Jobs/Assets • File Quality Score • Technical Steward • File Health Classification • File Health Description	• Contains Sensitive Data • Contains Restricted Data • Compliance Classification • Related Policies • Compliance Steward

Assets	Governance Dimension		
Fields	• Business Name (Title) • Business Description • Associated Metric(s) • Business Steward	• Technical Name • Field Data Type • Most Frequent Values • Most Frequent Patterns • Range of Values • Uniqueness • Field Quality Score • Field Quality Rules • Field Quality Thresholds • Technical Steward • Field Health Classification • Field Health Description	• Sensitive Indicator • Sensitive Classification • Associated Regulations • Required Risk Assessment • Related Risk Assessments • Compliance Classification • Related Policies • Compliance Steward • Last Compliance Review Date • Compliance Review Notes
Business Terms	• Business Name (title) • Abbreviation • Synonym(s) • Business Description • Associated Business Metrics(s) • Business Steward • Term Health Classification • Term Health Description • Certification Status		
Business Metrics	• Business Name (title) • Abbreviation • Synonym(s) • Business Description • Required Inputs & Dependencies • Aggregation Requirements • Calculation Example(s) • Associated Report(s) • Data Source Field(s) • Business Steward • Metric Health Classification • Metric Health Description • Certification Status		
Briefing Books	• Business Name (title) • Abbreviation • Synonym(s) • Business Description • Associated Report(s) • Associated Business Unit(s) • Owner Name Name • Business Steward		
Applications			• Technical Name • Business Name • Acronym • Synonyms • Description • Application Type • Primary Business Function • Associated Data Sources • Owner Name Name • Application Status • Application Location • Application Provider • Processing Frequency • Compliance Certifications • Compliance Steward

Assets	Governance Dimension		
Policies			• Policy Name • Abbreviation • Synonym(s) • Purpose • Scope and Applicability • Definitions • Roles and Responsibilities • Associated Policies • Additional information
Regulations			• Regulation Name • Abbreviation • Synonym(s) • Purpose • Scope and Applicability • Associated Policies • Additional information