

Holistic Data Governance: A Framework for Competitive Advantage



This document contains Confidential, Proprietary and Trade Secret Information (“Confidential Information”) of Informatica Corporation and may not be copied, distributed, duplicated, or otherwise reproduced in any manner without the prior written consent of Informatica.

While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Informatica does not accept responsibility for any kind of loss resulting from the use of information contained in this document. The information contained in this document is subject to change without notice.

The incorporation of the product attributes discussed in these materials into any release or upgrade of any Informatica software product—as well as the timing of any such release or upgrade—is at the sole discretion of Informatica.

Protected by one or more of the following U.S. Patents: 6,032,158; 5,794,246; 6,014,670; 6,339,775; 6,044,374; 6,208,990; 6,208,990; 6,850,947; 6,895,471; or by the following pending U.S. Patents: 09/644,280; 10/966,046; 10/727,700.

This edition published December 2012

Table of Contents

Executive Summary	2
Data Governance: Business Function, Not IT Project	3
Laying a Foundation for Data Governance Strategy	4
Understanding Levels of Data Governance Maturity	5
Developing a Data Governance Road Map	6
The 10 Facets of Data Governance	7
Vision and Business Case	7
People	8
Tools and Architecture	9
Policies	11
Organizational Alignment	13
Measurement	13
Change Management	16
Dependent Processes	16
Program Management	17
Defined Processes	17
Conclusion	19

Executive Summary

The goal of data governance is not just to clarify who “owns” data but also to optimize its value. The data itself is merely the means to the desired end of improved business performance. Accordingly, the responsibility for data governance efforts should fall at least as much on business as it does on IT—and preferably more.

This paper is intended to help data governance evangelists in both business and IT generate the momentum they need to make data governance an enterprise-wide priority. It provides a framework, tools, and vocabulary for educating business leaders about data governance, with the goal of influencing them to consider it a business issue deserving attention and investment.

Data Governance: Business Function, Not IT Project

Data governance often lands in IT's hands by default. When organizations focus on data for data's sake, though, they miss the broader picture: that data is only as valuable as the business processes, decisions, and interactions it enables and improves. The ultimate objective of data governance is to generate the greatest possible return on data assets. If business wants to be sure to capture critical opportunities to leverage data to support operations, strategy, and customer experience, it needs to govern data assets as it does other enterprise assets such as financial securities, cash, and human resources (HR).

Every organization, regardless of size or industry, has a core, self-sustaining finance function. In addition to its other operational responsibilities, finance is chartered to protect and optimize the business value of the organization's assets, specifically its financial assets. The CFO typically reports directly to the CEO and may even serve on the Board of Directors. And while many technologies enable the finance function—for example, ERP systems, spreadsheets, financial planning and analysis tools, and reporting tools—finance is not considered a technology or a technology market.

Most organizations also have an HR department. This is another core, self-sustaining business function chartered to protect and optimize the business value of the organization's assets, in this case its people, including employees and contractors. The head of HR is a senior business executive, often reporting directly to the CEO, or sometimes to the COO or CFO. HR relies for support on technology such as HR management systems, performance management applications, recruiting applications, and education enablement technologies. But as with finance, HR is not considered to be a technology or a technology market.

Data governance must fill the same sort of role as finance and HR: a coordinated enterprise effort that protects and optimizes the business value of the organization's assets—its data assets. As with other business functions, it requires people, policies, and processes with a clear way to measure success, compliance, and organizational effectiveness. Although data governance and its supporting data stewardship processes rely on technologies such as data integration, data quality, master data management, metadata management, data masking, data security, data archiving, and business intelligence software, data governance itself should not be considered a technology or a technology market.

Unfortunately, many business stakeholders conclude that because IT owns the databases, applications, and other systems that manage the data lifecycle, IT must be exclusively responsible for data governance as well. For the most part, data governance efforts are relegated to finite tactical IT projects such as data migrations or ERP upgrades that receive minimal or inconsistent grassroots business sponsorship and support. Unsurprisingly, these projects rarely scale to add strategic value across the organization.

The few organizations that do manage data governance as a business function are mostly found in highly regulated industries such as financial services. Even these organizations may manage the governance, risk, and compliance aspects of data governance as a self-sustained function, but they lack a holistic scope that extends to governing data to deliver business value beyond risk mitigation and compliance.

While managing data governance as a business function is not necessarily a requirement for success within your organization, it should be considered an aspirational goal, one that can lead to greater efficiencies, reduced costs, growing revenues, and better customer experiences.

Laying a Foundation for Data Governance Strategy

Although “data governance” and “data stewardship” are sometimes used interchangeably, they are two distinct concepts.

Informatica defines holistic data governance as:

A decision-making and cross-functional charter established to create data management policies, processes, and standards to optimize an organization’s return on its data assets.

Informatica defines holistic data stewardship as:

The front-line roles and responsibilities that ensure compliance with defined data governance guidelines to deliver trusted, secure data.

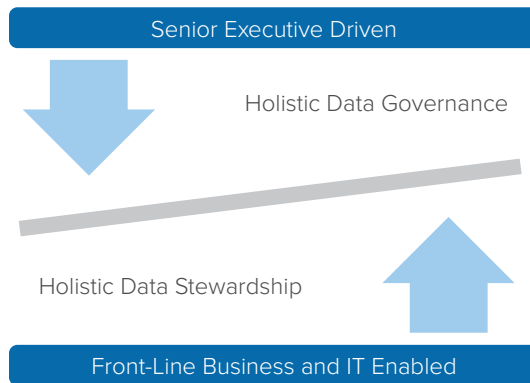


Figure 1. Clarifying Data Governance and Data Stewardship

In other words, data governance is a business function or program. As Figure 1 indicates, it is driven by top-down business leadership. It sets business priorities and objectives for maintaining trusted, secure data. It creates the policies, standards, roles, and responsibilities to ensure communication of the necessary dependent processes, owners, and measures of value.

By comparison, data stewardship involves people. It consists of the front-line business and IT stakeholders who put data governance into operational practice. They ensure ongoing compliance with and support for the stated policies and objectives.

Understanding this distinction is crucial; it clarifies the roles and responsibilities across business and IT that must be defined. While business sponsors and stakeholders focus their efforts on defining the business case, policies, and processes that will encompass the scope of the data governance effort, IT sponsors and stakeholders can simultaneously start to evaluate the current architecture and technology stack to ensure that IT has the necessary tools and investments to operationalize the supporting holistic data stewardship.

Understanding Levels of Data Governance Maturity

No organization begins to implement a data governance program from an entirely blank slate; every organization likely has some capabilities to leverage. Determining an organization’s current level of data governance maturity is a necessary first step in developing a customized plan that is both relevant and executable.

Informatica draws on Carnegie Mellon University’s popular Capability Maturity Model Integration (CMMI) process improvement approach to define six broad stages of maturity. (We added level 0 to account for the thankfully decreasing number of organizations that remain essentially unaware of a need for data governance; see Figure 2.) Four characteristics define the maturity level of an organization:

1. **Leadership.** From bottom-up grassroots efforts in Stage 1 to top-down executive-level leadership in Stage 5
2. **Scope.** From an ad hoc, siloed focus in Stage 1 to a self-sustaining core business function in Stage 5
3. **Measurement/Metrics.** From tactical technology metrics in Stage 1 to enterprise-wide, total impact to the business in Stage 5
4. **Data Governance Management.** From ad hoc exception-based management of the data governance effort itself in Stage 1 to management as a self-sustaining business function in Stage 5

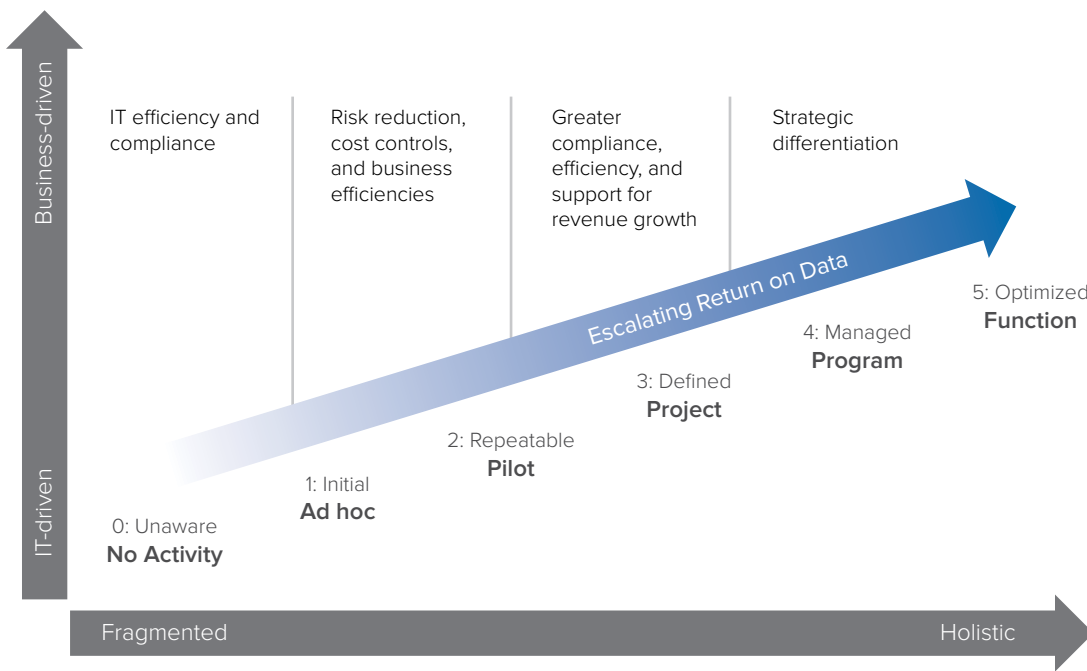


Figure 2. Stages of Data Governance Maturity

Entry-level Stage 1 maturity sees data governance as a fragmented silo-based initiative, while the aspirational goal of a self-sustaining data governance business function in Stage 5 sees it as a broad-based holistic enterprise program. (See Figure 3.) This natural evolution also moves data governance from being primarily an IT-driven project to a business-driven program. But these mapping of traits to specific stages of maturity are simply common trends. In reality, organizations can be at different levels of maturity while demonstrating different traits. For example, the data governance effort might be business-sponsored but defined as a project, rather than a program. This is a sign of room for improvement.

STAGE/TRAIT	0: UNAWARE	1: INITIAL	2: REPEATABLE	3: DEFINED	4: MANAGED	5: OPTIMIZED
Leadership	Minimal focus on data quality or security	Primarily grassroots driven by a few passionate individuals	Still grassroots but moving up to an EA or IT management level	Begins more top-down sponsorship, but primarily senior IT	Data governance program sponsored by business leaders	Top executive/board-level sponsorship and support
Scope	Data not prioritized in any meaningful, actionable way	Implement ad hoc rules, policies and/or standards as functional requirements into IT project	Documented IT governance and EA standards driving metadata reuse and improved collaboration across IT projects	Adopt competency centers and centers of excellence (e.g. ICC; BI CoE). IT-led, but business involved	Initiated as part of a broader strategic enterprise information management program	Data governance embraced as a self-sustaining core business function managing data as a corporate asset
Measurement	Zero measurement	Measured primarily on success of technology release	Measured primarily on improved IT efficiencies	Measured primarily on operational metrics and SLAs	Data governance lives through multi-phase, multi-year efforts but measured based on success of program	Measured on total impact to the business, not just confined to specific programs or strategies
Managed by	No activity	Ad hoc	Pilot	Project	Program	Function

Figure 3. Data Governance Maturity Stage Characteristics

Organizations can begin to achieve greater value from their data as early as Stage 1, but as Figure 2 clearly shows, greater progress escalates the return on data—that is, the degree to which organizations either reduce the cost of managing data or increase the overall value it delivers. In Stage 1, benefits generally involve IT efficiency and compliance. At higher levels of maturity, the benefits of data governance become increasingly strategic.

Developing a Data Governance Road Map

Determining the organizational level of maturity is a vital first step because it provides a clear picture of current strengths and areas for improvement. The process—which can be an internal effort—should include interviews of relevant business and IT staff, business risk surveys, business analyst time and activity analysis, and other techniques.

The goal is not just to establish a baseline but also to identify the level of data governance maturity the organization can reach. While Stage 5 is a worthy target, it may not be achievable in the near future given internal political, financial, and organizational realities. Instead, organizations should identify a goal level of maturity that can deliver business value while remaining realistically attainable and sustainable for three to five years.

The 10 Facets of Data Governance

Informatica's data governance framework provides context and understanding of the fundamentals necessary to build an effective data governance competency within an organization. It helps to identify strengths an organization can leverage for early value and momentum as well as areas for potential improvement and investment to mitigate organizational obstacles and risks.

Informatica's data governance framework is categorized into 10 complementary facets (see Figure 4):

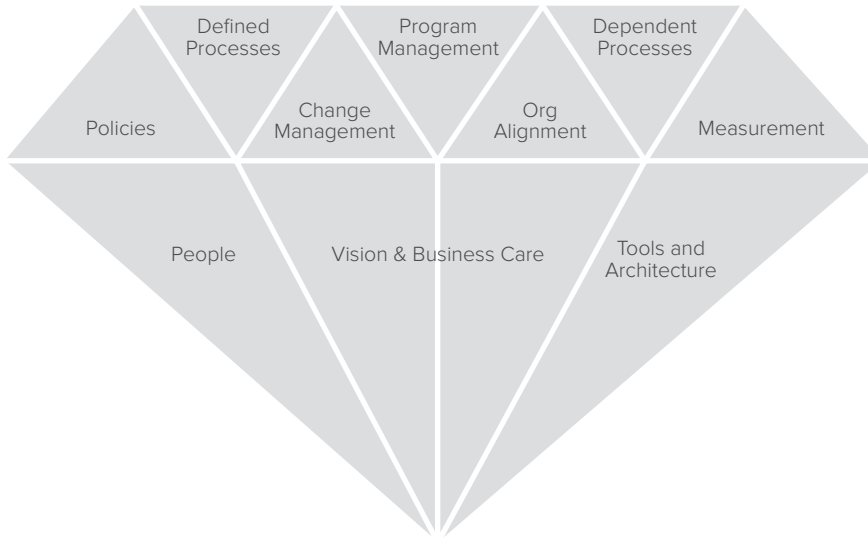


Figure 4. Informatica's Data Governance Framework

Vision and Business Case

The vision and business case must clearly articulate the business opportunity, both the broader strategic objective and the specific business opportunities on which to focus efforts. A vision statement is used to set an ultimate destination, but the business case must outline the journey to get there.

The vision defines the broader strategic objective.

Any strategic effort – not just data governance and information management—should establish and evangelize a foundational vision for all stakeholders with a clear definition of business goals and objectives. A vision statement should be actionable and specific, such as “Create an optimal customer experience by reducing the time to resolve support issues, deliver better targeted and relevant marketing, and ensure that sensitive customer data is protected and used appropriately.”

The business case identifies specific business opportunities.

The vision should look well into the future (three to five years minimum) in terms of the business value the data governance investments can deliver, but the business case must be pragmatic. Which prioritized business processes, decisions, or interactions will be influenced? How will actual value be delivered? For example, the vision statement above may lead to a quantitative business case focusing on just three targeted customer-facing processes: 1) productivity improvements through reduced average handle time in the call center's inbound support line, 2) revenue growth through increased campaign response rate and lower direct marketing costs, and 3) reduced enterprise risk through more effectively masked and secured customer information. Other organizations may choose to grow their data governance efforts one initiative at a time.

A well-rounded business case does not need to quantify the value of every potential dependent process, but it may qualitatively discuss other parts of the business that will benefit from these improvements, or mention future phase opportunities to be quantified later. Depending on executive budgeting approval requirements, you may only need to confirm that your “first phase” benefits will exceed your costs with returns in a “reasonable” amount of time. Success in the first phase will build the momentum for your phase two business case and beyond.

People

The right people are required to support, sponsor, steward, operationalize, and ultimately deliver a positive return on data assets. Organizations commonly form an executive steering committee to coordinate communication, prioritization, funding, conflict resolution, and decision making across the enterprise. With the right executive sponsorship, a steering committee can be created early, but grassroots efforts might require some momentum before senior leaders are willing to commit their support.

An effective data governance program should include all of these roles:

- **Executive sponsor.** The optimal executive sponsor(s) will be a CxO-level executive whose responsibilities span functional, line-of-business, application, and geographic silos. The earlier that sponsors are identified, the better, because they drive resource allocation, staffing, funding, business prioritization, and cross-functional collaboration. To be effective, a sponsor must be an active participant and evangelist.
- **Data steward/data quality steward.** Data stewards are the business and IT subject matter experts who can most effectively translate how data and systems influence the business processes, decisions, and interactions most relevant to the organization. The business stewards must be IT-savvy; the IT stewards must be business-savvy. Both must be strong communicators and facilitators across the part of the organization they represent. Experienced business analysts with the expertise to bridge business and IT communications often make the best business stewards, while data and enterprise architects and senior business systems analysts can provide critical perspectives as IT stewards.
- **Data governance leader.** The primary responsibility—and skill—of the data governance leader is not to care who “wins.” The data governance leader coordinates tasks for data stewards, helps communicate decisions made by stewards to relevant stakeholders, drives ongoing data auditing and metrics that assess program success and ROI, and is the primary point of escalation to the executive sponsor and steering committee. A program or project management office (PMO) may sometimes serve in this role.

Tools and Architecture

As Figure 5 indicates, enterprise and data architects should consider the full lifecycle of critical enterprise data. This includes:

- Upstream on-premises transactional/operational applications, systems, and processes that create, update, import, or purchase data.
- Downstream on-premises analytical applications, systems, and processes that consolidate, reconcile, deliver, and consume data.
- Growth of off-premises sources and targets of data, including cloud-based applications and platforms, social data, mobile devices, third-party data feeds, sensor data, and Hadoop analytic environments.
- Supporting data management infrastructure investments that enable and ensure compliance with the organization’s unique requirements for delivery of “the right data at the right time with the right latency of the right quality and security in the right context.”
- Assessment and delivery of the shared capabilities that must be made available across the enterprise data architecture, not confined within specific applications or tools. A common approach includes an Integration Competency Center with responsibility for a unified data management platform.

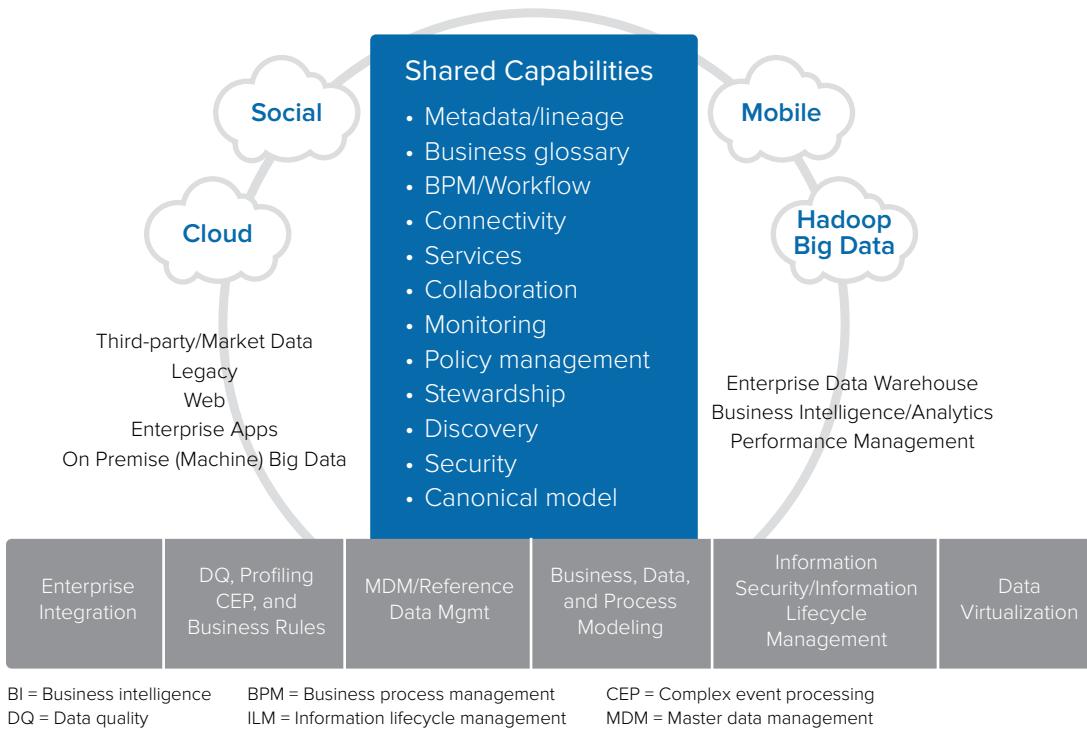


Figure 5. Architectural Scope of Data Governance

Specific enabling software capabilities that should be considered to help launch a data governance effort include:

- **Data profiling.** Data profiling software helps business analysts and data stewards answer the questions “What does data look like today?”, “How does data in one system relate to data in another system?”, and “What rules and policies should we consider defining to improve?”
- **Data discovery.** While data profiling allows in-depth analysis of specified data sets, data discovery allows organizations to identify where any data anomaly or business scenario occurs across all data sources. For example, the data privacy organization may require the ability to identify where personally identifiable information (PII) is used and how that relates to specific business rules/processes where obfuscation or data masking needs to be used.
- **Business glossary.** A business glossary allows business and IT stewards to capture and share the full business context around critical data. In addition to the expected definitions of core data entities and attributes, context can also include rules, policies, reference data, free-form annotation, links, and data owners. Some organizations simply manage their shared definitions in Word documents or spreadsheets, an approach that typically focuses on the terms and definitions but misses the broader context. A packaged business glossary enables collaboration across the business and IT roles that create, approve, and consume these definitions. This minimizes the risk of redundancy, definition stagnation, and versioning conflicts. A good business glossary should also provide access capabilities that enable sharing of the business vocabulary across the organizations. A common business vocabulary that is not available to everybody all the time has limited value.
- **Metadata management/data lineage.** The ability to reconcile and provide transparency and visibility to the supporting metadata of critical data is a foundational element of a data management reference architecture. Data lineage visualization and auditing capabilities allow data architects and stewards to visualize the flow of data through their environment and effectively assess impact analysis of potential changes to data definitions, rules, or schemas, as well as root cause analysis capabilities when responding to a data quality or security failure. IT staff ranging from data modelers to enterprise architects, business systems analysts, developers, and DBAs often manage the technical metadata, while business analysts and business stewards are often responsible for the business-oriented metadata. Ideally, the business glossary should be well-integrated with the metadata solution to enable an unambiguous collaboration between business and IT.

In addition to these four capabilities, architects should also consider where and how they want to manage their data modeling, process modeling, data quality, data privacy, master data, data monitoring and auditing, workflow management, and collaboration capabilities. In addition, they must determine how data stewards will be notified and how these stewards should mitigate exceptions to any established data quality, data retention, or privacy rules.

Policies

Business policies and standards are critical for any data governance function. Common policies that must be agreed upon, documented, and complied with include data accountability and ownership, organizational roles and responsibilities, data capture and validation standards, information security and data privacy guidelines, data access and usage, data retention, data masking, and archiving policies. Common challenges in defining these policies often include the following three factors:

- **Analysis paralysis.** Policies, by definition, set parameters on how employees and other stakeholders are expected to behave. People resist change, especially if they're the ones affected and the impacts are uncertain. An executive sponsor or steering committee can use its influence to help keep debate constructive and prevent politics from stalling efforts.
- **Noncompliance.** Don't waste time defining policies that won't be used until there is top-down support to both provide an incentive and enforce compliance.
- **Usability.** It's one thing to have executive support to ensure compliance; it's quite another to understand how to ensure compliance. If the policy isn't written in a way that can be clearly understood as requirements to be consumed by business process, application, and data management stakeholders, how can they deliver the necessary process, system, and rule changes necessary to comply? For example, translating a complex record retention schedule into simplified, higher-level categories is now considered a best practice.

Many organizations have existing governance programs that they can leverage to launch data governance. For example, an organization may document business- or IT-driven policies that set parameters on how enterprise data should be managed and used, although this documentation effort may not be labeled data governance. IT governance, enterprise architecture (EA) standards, and any number of IT-led competency centers or centers of excellence may also offer defined standards and policies on how best to capture, update, and share relevant data and metadata across the organization. The organization's Chief Risk Officer and CFO may have issued governance, risk, and compliance policies, information security or data privacy policies, and any number of policies to ensure compliance with external government and industry regulations. Scoping and defining policies to document and implement can begin by reconciling which existing policies should be owned by the data governance effort, which should simply be recognized and complied with, and which should be replaced.

A data governance initiative will be responsible for documenting and maintaining policies such as these:

- **Data accountability and ownership.** These policies spell out which senior business leaders or groups (e.g., a steering committee) are accountable for the quality and security of critical data. The policy must outline what ownership actually means, define the rights and responsibilities of the owners, and indicate whether and how those responsibilities change over time.
- **Organizational roles and responsibilities.** These policies document and make clear the responsibilities of business and IT data stewards, the data governance leader, and other dependent stakeholders.
- **Data capture and validation standards.** These policies define minimum required data capture standards, data validation rules, and reference data rules, among others. The goal is to ensure that the people, processes, and systems that capture, import, update, transform, or purchase critical data do so in a consistent, standardized manner with a focus on quality that ensures fitness for enterprise use. Best practices also dictate that data retention should be defined when data gets captured.

- **Data access and usage.** Data usage policies ensure appropriate use of data by appropriate stakeholders. Limiting access to sensitive or confidential information ensures compliance with edicts such as insider trading regulations and the Payment Card Industry Data Security Standard (PCI-DSS). Usage policies extend beyond regulatory compliance to ensure optimal use of data assets. For example, contact management policies are used to coordinate, prioritize, and minimize multichannel customer communications across sales, marketing, and service organizations. These policies help to reduce lower value contacts and avoid the risk of customers feeling spammed—which may lead them to opt out of future communication.
- **Arbitration and adjudication.** These policies are defined to ensure a formal process is in place to address and resolve data-related conflicts, particularly how to address conflicting points of view on the proper definition and usage of important enterprise data.

The risk, security, and compliance stakeholders who own and monitor compliance with information security and data privacy policies must be active contributors to the data governance program. They oversee policies such as these:

- **Customer communication privacy preferences.** These privacy policies ensure compliance with anti-spam legislation, “do not call” registries, and customer marketing contact management best practices. These policies are meant to be transparent, customer-facing documents that should clarify whether the organization has adopted an opt-in policy (will never send marketing communications unless customer expressly permits) or opt-out policy (will always send marketing communications unless customer expressly requests a stop). These policies can go as granular as necessary and can request opt-in/opt-out at the marketing communication channel level (e.g., phone, mail, email, text, social) and even ask customers to set preferences for which products and/or services they are interested in learning more about.
- **Data masking.** These data privacy policies define and classify sensitive data, identify where it resides, and clarify when it needs to be encrypted appropriately and consistently across multiple applications and database instances. Data masking policies are critical to ensure compliance with mandated data protection rules and standards such as PCI-DSS, personal health information (PHI), and PII. Organizations demand the same data security policies for archiving as they do for their testing and even production environments, so the data governance effort must include all of these instances in its scope when managing data privacy. Data security is a common classification as well, enabling the company to define top secret, company confidential, and public data and maintain an aligned data handling policy.
- **Data archives and test data subsets.** Organizations focus on these procedures to dramatically reduce the inactive data in production and legacy environments in order to improve system performance and to cut down on infrastructure (i.e., server and storage) costs. These policies, when effectively implemented, should also reduce data storage costs in development and test environments.
- **Data retention.** Retention policies must align with a corporate records retention schedule (RRS), meet the organization’s need for specific data, and balance the desire to archive and purge unused data to reduce storage costs and business risk with the need to comply with business and legal discovery retention management requirements. These policies will clarify what data needs to be stored for how long, in what format, applying what rules, with what level of masking or encryption, and with what access guidelines.

Organizational Alignment

Organizational alignment focuses on the working relationships among roles, addressing questions such as these:

- Who will be the executive sponsor for the organization's data governance efforts?
- Will there be an executive steering committee?
- Who are the business data owners?
- What are the escalation paths for policy and data conflicts?
- Is data steward a full-time or a part-time role?
- Do data stewards hold solid-line or dotted-line reporting relationships to the executive sponsors?

We recommend using a responsibility assignment matrix such as RACI (RACI defines roles of responsible, accountable, consulted, and informed) or DACI (DACI defines the roles of driver, approver, contributors, and informed) to help align and set expectations with the various stakeholders involved in all aspects of the data governance effort.

Measurement

Data governance must be measured at three distinct levels. First, at the program level, the organization must identify and highlight the qualitative level of organizational influence and impact the data governance efforts deliver. Next, stewards need operational data monitoring to evaluate how the data is behaving against expected policy and validation baselines. Last, and most importantly, sustaining data governance momentum requires quantitative business value measurement that links data management efforts to real business value such as revenue growth, cost savings, risk reduction, efficiency improvements, and customer satisfaction.

We recommend building the measurement strategy as follows:

- **Start with data governance program effectiveness to satisfy sponsors.** Early in the life of a data governance initiative, the biggest struggle is often getting business and IT stakeholders to pay attention. An important measure of success is the program's level of engagement, participation, and influence. How many lines of business, functional areas, system areas, project teams, and other parts of the organization have committed stewardship resources or sponsorship? In addition, categorize and track status of all issues that come in to the data governance function, and capture all other types of value-added interactions such as training, consulting, and project implementation support. While these metrics may not demonstrate business value, they will help early stage data governance efforts demonstrate progress in operationalizing data management efforts.

- **Develop operational data quality and policy auditing metrics to focus data stewards.** The business case and ROI are not about the data, but the resulting business rules, policies, processes, and standards are. Business and IT data stewards alike are responsible for monitoring data to ensure compliance with these standards and to mitigate data quality, privacy, and security issues as needed. To do so, they need predefined stewardship workflows that include operational metrics such as data accuracy, completeness, integrity, uniqueness, consistency, and standardization. They also need to be able to conduct audits ensuring compliance with contractual service-level agreements (SLAs), data retention, and privacy and security policies. Stewards' visibility into these metrics must include metadata and data lineage analysis capabilities for root cause and impact analysis as well as transparency and audit trails for compliance purposes. In addition, managers should track data stewards' own productivity via metrics on how many assigned issues have been resolved, how many records have been cleared, and how many remain in their queue.
- **Model business value and ROI measures to maintain momentum with business leadership.** This step is the critical factor elevating data governance from a one-off IT project to a sustainable aspect of doing business. Business value from data governance investments can include a variety of benefits such as minimizing penalties for regulatory noncompliance, reducing enterprise risk, lowering costs, optimizing spending, improving operational efficiencies, increasing top-line revenue growth, and optimizing customer experience and satisfaction.

Building a business case with real business value drivers is one of the most common challenges for those evangelizing data governance efforts within their organization. Kicking off or piloting a data governance initiative by attaching it to a prioritized and funded business project is a proven way to limit scope and focus the value on the goals of that project. Each organization will find its own particular combination of business drivers compelling; the examples in the following chart are meant to inspire rather than prescribe.

EXAMPLES OF COMMON DATA GOVERNANCE BUSINESS DRIVERS				
Growth	Revenue growth: <ul style="list-style-type: none"> • Increase cross-sell and upsell opportunities • Increase marketing campaign response rates • Improve account management 	Market share growth: <ul style="list-style-type: none"> • Improve competitive insights • Enhance account qualification 	Footprint growth: <ul style="list-style-type: none"> • Support regional expansion • Extend vertical industry expansion • Advance multichannel distribution efforts 	Product portfolio growth: <ul style="list-style-type: none"> • Support M&A, OEM, and partnering strategies • Enter new markets • Depart dying markets
Compliance & Risk Reduction	Compliance risk: <ul style="list-style-type: none"> • Enable transparency for financial reporting and auditability (SEC, SOX, Basel II) • Improve readiness for an e-discovery event • Adhere to corporate data security policies and standards 	Information security/privacy: <ul style="list-style-type: none"> • Secure or mask sensitive, classified and confidential information • Ensure privacy compliance for marketing communication preferences • Enable PCI compliance 	Avoiding and defense from illegal activities: <ul style="list-style-type: none"> • Comply with anti-money laundering (AML) and Patriot acts • Comply with U.S. Department of State country policies and embargoes • Enhance cyber-security to protect from virtual attacks 	Contractual risk: <ul style="list-style-type: none"> • Reduce counterparty risk
Efficiency & Cost Reduction	Cost savings: <ul style="list-style-type: none"> • Reduce cost to store duplicate or low value data • Cut direct marketing spend on postage and collateral to invalid or duplicate customers 	Reduced/optimized spending: <ul style="list-style-type: none"> • Enable spend management analysis to ensure optimal procurement pricing 	Improved efficiencies: <ul style="list-style-type: none"> • Improve decision support process by reducing time needed to collect and cleanse data • Reduce call center response times with access to timely, trusted customer and product information 	
Strategic Differentiation	Optimal customer experience: <ul style="list-style-type: none"> • Enable know-your-customer and other customer intelligence functions • Increase customer loyalty and satisfaction 	Optimized supply chain: <ul style="list-style-type: none"> • Centralize purchasing to reduce redundancy that can lead to excess inventory • Reduce risks of stock-out costs while minimizing excess inventory 	Business transformation: <ul style="list-style-type: none"> • Move from single channel to multichannel business model • Change from physical to digital products • Support mergers or divestitures of core businesses 	

Some business opportunities are more obviously quantifiable than others. However, if the data governance team can provide a balanced report that includes at least some transparent quantitative business value on certain targeted processes, the senior leadership at most organizations will accept more qualitative benefits for others, given enough anecdotal support to demonstrate how key influencers believe data governance efforts are helping. In the end, the ROI should at minimum pay for resources that are being dedicated to the project.

Change Management

No matter how compelling the vision and business case, making data a true corporate asset is a major culture shift for most organizations. Accomplishing this shift will likely require significant behavioral change across the workforce (and possibly in the partner/supplier ecosystems) to create an organizational culture that values data properly. Support for these organizational, business process, and policy changes must include training, communication, and education, with a “carrot and stick” performance management program to incentivize good data practices while discouraging past undesirable behaviors. Data governance programs must also include time, resources, and a commitment from management to invest in necessary change management.

Dependent Processes

An organization cannot fully govern its critical enterprise data until it understands the entire lifecycle of the data. Downstream operational and analytical processes that consume and derive insight and value from data are usually top of mind and often form the basis of the business case for data governance. However, to deliver trusted, secure data, organizations must also understand the upstream business processes that create, update, transform, enrich, purchase, or import data. These upstream processes often create a significant percentage of the “garbage in/garbage out” problems that the data governance effort is chartered to resolve.

Dependent processes are categorized into three broad areas: upstream processes, stewardship processes, and downstream processes. (See Figure 6.) An effective data governance organization will accept responsibility for assessing and improving all of the processes that touch the data and influence its usability.

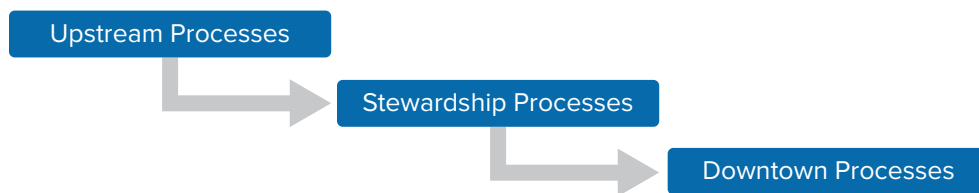


Figure 6. Data Governance Responsible For All Data Lifecycle Processes

Upstream processes. These are the business processes that capture, create, import, purchase, transform, or update data and introduce it into the organization’s information ecosystem. One of the most common, and most intractable, data governance challenges comes from the fact that the people responsible for these upstream processes rarely have visibility into—or incentive to care about—who is consuming this data downstream and why. A senior and influential executive sponsor plays a vital role in evangelizing and enforcing the data capture and maintenance policies that the data governance organization generates.

Stewardship processes. In the lifecycle of physical data itself, the stewardship process stage involves applying the data policies, business rules, standards and definitions created as part of the data governance program. The automated application of these rules in a system-centric workflow may appear as service-enabled or application-specific rules that archive, cleanse, enrich, mask, match, merge, reconcile, repair, validate, verify, or otherwise improve the security and quality of data. In a more human-centric workflow, stewardship processes facilitate the manual identification, notification, escalation, and mitigation of exceptions to the automated rules and policies.

Downstream processes. These are the operational and analytical processes that consume, protect, archive, purge, and otherwise extract insight and value from data. Delivering significant business value and ROI against these downstream processes is the way to persuade executive sponsors to support changes to upstream processes, systems, and organizational behaviors—and to allow investment in the organizational and technology improvements that enable the stewardship processes. This is why they generally form the foundation of the vision and business case for the data governance program.

Program Management

A multiphase, multiyear plan for starting small and growing into cross-enterprise, self-sustaining holistic data governance doesn't manage itself. Whether through an official program management office (PMO) or a team of program drivers, data governance efforts need skilled project/program management professionals to coordinate the complex interactions, communications, facilitations, education, training, and measurement strategy. Effective program management can ensure adoption, visibility, and momentum for future improvements.

It is important to recognize the significant amount of coordination, facilitation, and communication necessary to evangelize, prioritize, measure, and evolve data governance from a pilot project to a foundational way of doing business. Ultimately, this must be someone's full-time job. Although the business must own and accept accountability for data governance, and the resulting policies, rules, and standards, the most significant leadership roles will likely end up with IT, which typically offers the strongest program/project management skills within an organization.

Defined Processes

If data governance is a business function like finance, and finance is made up of processes such as accounts payable, accounts receivable, payroll, and financial planning, it would follow suit that data governance has its own set of processes. Figure 7 illustrates the stewardship processes that make up the data governance function. Some examples include the processes that cleanse, repair, mask, secure, reconcile, escalate, and approve data discrepancies, policies, and standards. We have segmented the many processes of data governance into core stages: discover, define, apply (rules and policies), and measure/monitor:

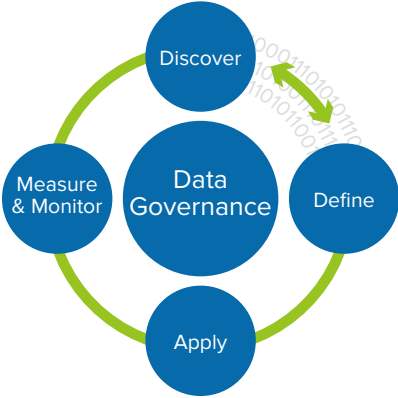


Figure 7. The Process Stages of Data Governance

- **Discover** processes capture the current state of an organization's data lifecycle, dependent business processes, and supporting organizational and technical capabilities, as well as the state of the data itself. Insights derived from these steps define the data governance strategy, priorities, business case, policies, standards, architecture, and the ultimate future state vision.

This stage includes the supporting processes of data discovery, data profiling, inventory of current state data and processes, CRUD (create, read, update, delete) analysis across the data lifecycle, and assessments of organizational, people, and technology capabilities. This process runs parallel to and is iterative to the define process stage: Discovery drives definition, while definition creates a more targeted focus for discovery.

- **Define** processes document data definitions and business context associated with business terminology, taxonomies, and relationships, as well as the policies, rules, standards, processes, and measurement strategy that must be defined to operationalize data governance efforts.

This stage includes the supporting processes of business glossary creation, data classification, data relationship, hierarchy and reference data definitions, and the definition of supporting business rules, policies, and key performance indicators (KPIs). This process runs parallel to and is iterative to the discover process stage mentioned above.

- **Apply** processes aim to operationalize and ensure compliance with all the data governance policies, business rules, stewardship processes, workflows, and cross-functional roles and responsibilities captured through the discover and define process stages.

This stage includes the processes enabling automation of the business rules and policies defined within applications, processes, and services. It also includes the processes that operationalize the supporting human-centric business and IT workflows that manually mitigate exceptions to data quality or data security rules and policies.

- **Measure and monitor** processes capture and measure the effectiveness and value generated from data governance and stewardship efforts, track compliance with and exceptions to defined policies and rules, and ensure that data assets and their lifecycle are transparent and auditable.

This stage includes supporting processes such as proactive data quality and policy compliance monitoring, reactive auditing of operational data quality, and data lineage analysis for root cause and impact assessment. In addition, the stage includes the processes measuring the data governance program's quantitative and qualitative effectiveness and the ongoing business value delivered.

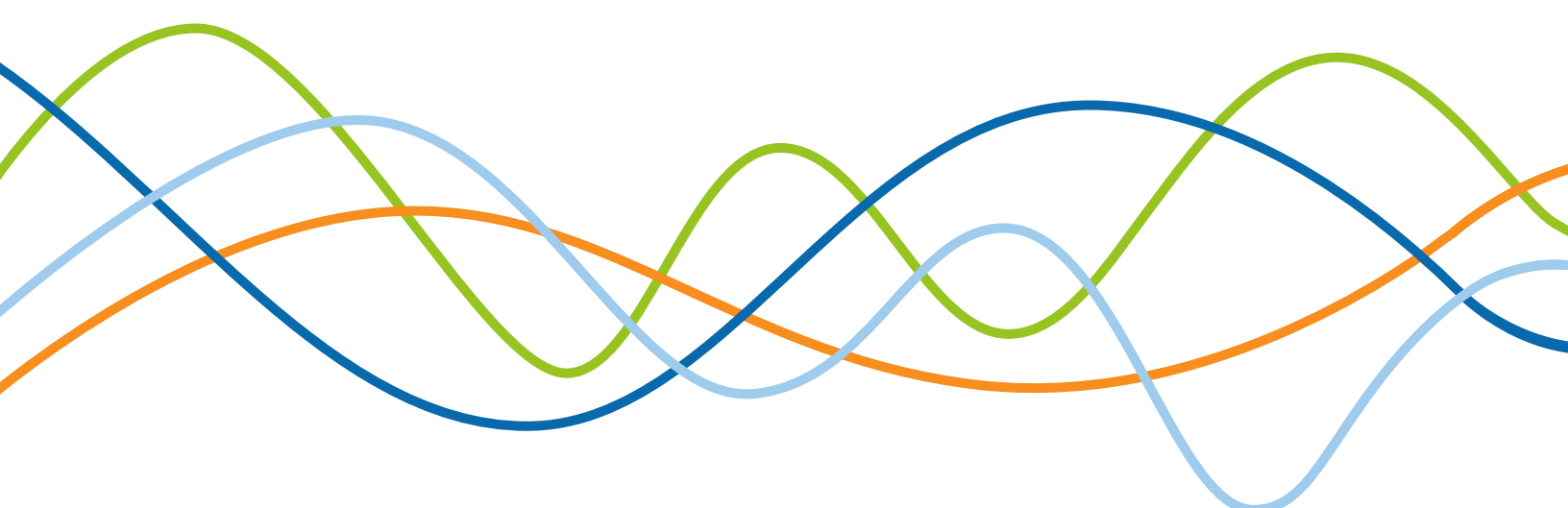
Conclusion

An organization addressing targeted data quality concerns within a single application will demand less organizational and technical complexity from its data governance efforts than an organization trying to master multiple data domains across a large global enterprise, which requires high levels of coordination and formal processes. Nevertheless, both types of organizations can benefit from Informatica's multifaceted framework for launching or expanding a holistic data governance program—as long as they keep in mind that data governance is not about data for its own sake.

Data governance is not an IT project, but a business imperative. Data governance managers and evangelists must focus on this fact and create a business case that clearly demonstrates how leveraging data enables people, process, organization, and technology improvements for greater business value. Only in this way can they persuade executive leaders to make ongoing investments in data governance and the data management technologies that enable it.

ABOUT INFORMATICA

Informatica Corporation (NASDAQ: INFA) is the world's number one independent provider of data integration software. Organizations around the world rely on Informatica for maximizing return on data to drive their top business imperatives. Worldwide, over 4,630 enterprises depend on Informatica to fully leverage their information assets residing on-premise, in the Cloud and across social networks.



INFORMATICA[®]

Worldwide Headquarters, 100 Cardinal Way, Redwood City, CA 94063, USA
phone: 650.385.5000 fax: 650.385.5500 toll-free in the US: 1.800.653.3871
informatica.com [linkedin.com/company/informatica](https://www.linkedin.com/company/informatica) twitter.com/InformaticaCorp

© 2012 Informatica Corporation. All rights reserved. Printed in the U.S.A. Informatica, the Informatica logo, and The Data Integration Company are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

IN09_1212_02297