

**TDWI**

MONOGRAPH SERIES

JULY 2008

# The Four Imperatives of Data Governance Maturity

By Philip Russom  
Senior Manager, TDWI Research  
The Data Warehousing Institute

SPONSORED BY



## Table of Contents

|  |    |
|--|----|
| Defining Data Governance and Its Maturation Cycles .....                   | 3  |
| The Four Imperatives of Data Governance.....                               | 3  |
| The TDWI Maturity Model .....  | 4  |
| DG Imperative 1: Maintain a Cross-Functional Team and Process .....        | 5  |
| DG Imperative 2: Align with Data-Intense Business Initiatives .....        | 7  |
| DG Imperative 3: Govern Data Usage via Technical Implementations.....      | 9  |
| DG Imperative 4: Automate DG Processes via Technical Implementations ..... | 10 |
| The TDWI Maturity Model for Data Governance .....                          | 14 |
| Recommendations.....   | 15 |

## About the Author



**PHILIP RUSSOM** is the senior manager of TDWI Research at The Data Warehousing Institute (TDWI), where he oversees many of TDWI’s research-oriented publications, services, awards, and events. Before joining TDWI in 2005, Russom was an industry analyst covering BI at Forrester Research, Giga Information Group, and Hurwitz Group. He’s also run his own business as an independent industry analyst and BI consultant, and was contributing editor with *Intelligent Enterprise* and *DM Review* magazines. Before that, Russom worked in technical and marketing positions for various database vendors. You can reach him at [prussom@tdwi.org](mailto:prussom@tdwi.org).

## About Our Sponsor

DataFlux enables organizations to analyze, improve, and control their data through an integrated technology platform. With DataFlux enterprise data quality and data integration products, organizations can more effectively and efficiently build a unified view of customers, products, suppliers or any other corporate data asset. DataFlux is a wholly owned subsidiary of SAS. DataFlux customers can rapidly assess and improve problematic data, building the foundation for enterprise data governance. Effective data governance delivers high-quality information that can fuel successful enterprise efforts such as risk management, operational efficiency, and master data management (MDM). To learn more about DataFlux, visit [www.dataflux.com](http://www.dataflux.com).

# Defining Data Governance and Its Maturation Cycles

Anytime data crosses an organizational boundary, it should be governed, whether you’re sharing data among business units internally or publishing data to customers, partners, auditors, and regulatory bodies externally. Furthermore, we now live in the “age of accountability,” which (among other things) demands stricter oversight for data usage, quality, privacy, and security. User organizations are under renewed pressure to ensure that compliance and accountability requirements are met as the scope of data integration broadens. In response to this situation, many organizations are turning to data governance.

TDWI’s definition of data governance covers most of its components and goals:

*Data governance (DG) is usually manifested as an executive-level data governance board, committee, or other organizational structure that creates and enforces policies and procedures for the business use and technical management of data across the entire organization. Common goals of data governance are to improve data’s quality; remediate its inconsistencies; share it broadly; leverage its aggregate for competitive advantage; manage change relative to data usage; and comply with internal and external regulations and standards for data usage. In a nutshell, data governance is an organizational structure that oversees the broad use and usability of data as an enterprise asset.<sup>1</sup>*

## The Four Imperatives of Data Governance

Data governance tasks distill down to four broad imperatives.

As you can see, there’s a lot to data governance. Luckily, it’s not as difficult to grasp as it seems, because the many goals and tasks associated with DG distill down to four imperatives, which in turn group into a pair of organizational imperatives and a pair of technical ones (see Figure 1).

|                                   |  |
|-----------------------------------|--|
| <b>Organizational Imperatives</b> | <ol style="list-style-type: none"> <li>1. Maintain a Cross-Functional Team and Process</li> <li>2. Align with Data-Intense Business Initiatives</li> </ol>         |
| <b>Technical Imperatives</b>      | <ol style="list-style-type: none"> <li>3. Govern Data Usage via Technical Implementations</li> <li>4. Automate DG Process via Technical Implementations</li> </ol> |

Figure 1. The Four Imperatives of Data Governance, in pairs for organizational and technical tasks

The four imperatives of DG have characteristics relative to maturation:

**Each DG imperative has lifecycle stages.** These unfold over time, and progression through the stages amounts to a form of maturation. The next section of this Monograph defines the TDWI Maturity Model, which consists of such lifecycle stages arranged in a maturation sequence.

**The imperatives as a group imply a time sequence.** For example, it’s obvious that imperative 1 must create a cross-functional team before imperative 2 can align team goals with business initiatives. Less obvious is that imperative 3 should be governing IT systems before imperative 4 starts using IT systems to automate governance processes. Although dependencies like these determine an order for commencing the imperatives, the imperatives must eventually coexist and interact.

<sup>1</sup> For a detailed discussion of data governance, see the TDWI Best Practices Report *Data Governance Strategies*, available online at [www.tdwi.org/research/reportseries](http://www.tdwi.org/research/reportseries).

**Maturation can force an imperative to restart.** For example, many DG teams must reorganize when they consolidate or coordinate with other governance teams. Imperative 2 is inherently iterative, because business strategy and external requirements (like compliance) change periodically. Likewise, imperatives 3 and 4 are ruthlessly iterative, since most of DG maturation involves incrementally taking on new systems to govern and new tools for greater DG automation.

**Each DG imperative is a critical success factor.** For a data governance program to survive over time and to scale to broad influence, it needs all four imperatives matured into advanced stages. That’s why they’re called *imperatives*.

### The TDWI Maturity Model

The lifecycle stages of a DG imperative—or a complete DG program—can be described by the TDWI Maturity Model. Let’s review the model before plotting DG’s maturation cycles on it.

TDWI’s Maturity Model is a generalized method for plotting the progress of any business or technical solution.

Figure 2 illustrates the TDWI Maturity Model, which arranges generalized lifecycle stages in a left-to-right maturation process on the X axis. On the Y axis, a traditional bell curve represents the rough percentage of organizations in a given maturation stage. Since the DG programs of most organizations are today stuck in the child and teenager stages, the bell is in the middle of the chart. The bell should move to the right over time, as more organizations move into advanced stages.

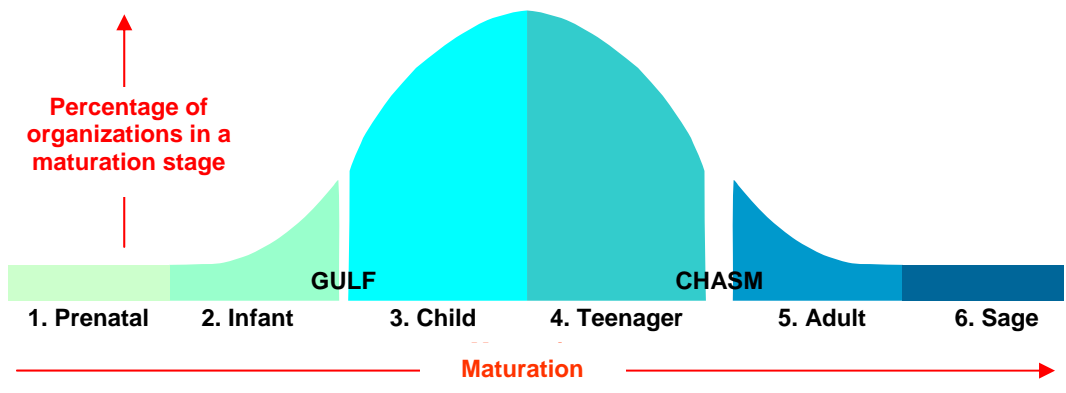


Figure 2. The TDWI Maturity Model

TDWI’s maturity model includes six lifecycle stages (which group into pairs), plus two gaps representing barriers commonly encountered as an organization progresses through the stages. Here’s how the stages apply to most business initiatives or technical implementations.

- **Prenatal and Infant Stages.** In the prenatal stage, an organization relies on manual means applied in an ad hoc manner as a low-end solution to a business or technology problem. By the infant stage, a study of requirements has led the organization to a specific technology or practice, and they have initiated proof-of-concept, prototype, and phase 1 solutions.
  - **The Gulf.** Vaulting this hurdle depends on the organization institutionalizing the solution concepts it has already proved.
- **Child and Teenager Stages.** The child stage is an exciting time of growth and learning, as the organization expands the new technology or practice it recently committed to. But growth often occurs in a limited context, such as a handful of departments or shortlist of IT systems, so growth slows down in the teenager stage.
  - **The Chasm.** More daunting than crossing the gulf, crossing the chasm successfully involves dramatic global changes, like enterprise adoption or solution re-architecture.

- Adult and Sage Stages.** Solution best practices and technology implementation details developed in the teenager stage continue to mature in the adult stage. The Big Picture demanded on the right side of the chasm leads to maturation in areas like cross-departmental coordination and technology scalability. The silo deployments of the child stage are gone, replaced in the sage stage by centralized organizational control and technology integration.

TDWI's Maturity Model helps organizations plan, assess, and revitalize initiatives.

The point of the TDWI Maturity Model is to chart a course for organizations that need to know where to start and where to go with a particular type of initiative. Although each organization must adapt it to its own situation, the maturity model assists with planning and provides an objective yardstick for assessing the current state of an initiative. The model also gets business and technical people brainstorming about improvements and additions that can revitalize existing initiatives.

This Monograph charts the four DG imperatives on the TDWI Maturity Model one at a time to show each one's dynamics. The Monograph concludes by pulling the four together into a unified maturity model to illustrate where the imperatives align and interact—and where they don't. The point is to put the common starting points and milestones of a data governance program into a linear context, so organizations know where to start, where to go, and which pitfalls to avoid.

## DG Imperative 1: Maintain a Cross-Functional Team and Process

The first imperative of data governance—*maintain a cross-functional team and process*—is very much about organizational dynamics on different levels. On one level, people first coalesce their DG efforts into a permanent organizational structure, like a DG committee. On a grander level, this relatively small committee must progressively tame the larger organization, both convincing and coercing it to comply with the committee's growing list of policies and procedures for data access and usage. (See Figure 3.)

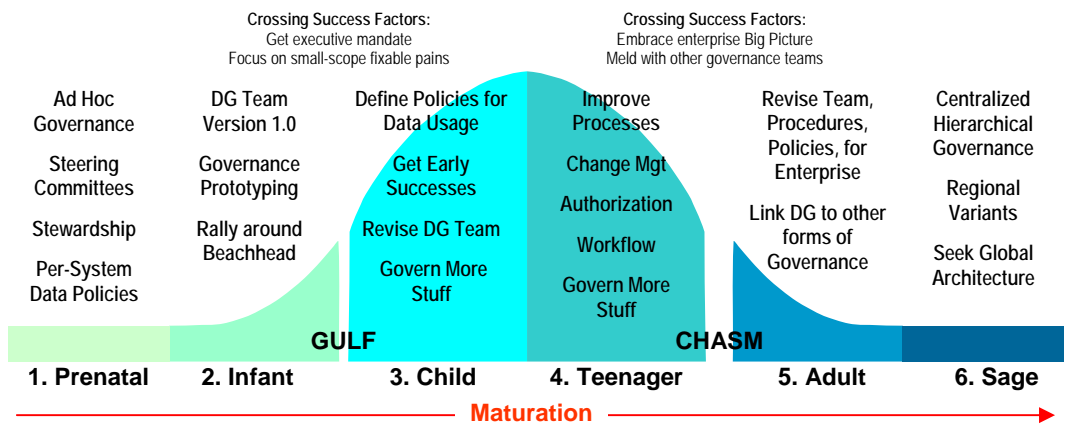


Figure 3. Maturity Model for DG Imperative 1: Maintain a Cross-Functional Team and Process

In the prenatal lifecycle stage, controls on data usage (which resemble data governance) are occasional and informal. For example, a line of business (LOB) manager may train his/her people to use departmental applications in a particular manner. Likewise, some technical developers may work with managers to devise user roles and security procedures for applications. These efforts, though well intentioned, tend to follow a local interpretation of a global directive, and so may or may not comply with corporate or external regulations. Consistency for data access and usage across departments and applications is unlikely.

The infant lifecycle stage commences as people from diverse functions come together to increase compliance or consistency for data usage. The result is some kind of organizational structure,

typically a so-called data governance committee or board. The people on DG committees often draw from their experiences with similar cross-functional organizations, like data stewardship programs, data warehousing teams, steering committees, advisory boards, and competency centers. The resulting membership of a DG committee is diverse, involving a mix of business and technical people from multiple departments and business units. This diverse lot is held together, in the beginning, by common pains resulting from a lack of controls on data and its usage.

Once formed, the data governance committee must cross the gulf to the child stage by governing the data usage of specific initiatives and implementations. This is a critical moment that depends on support from an attentive executive sponsor who can articulate the DG vision and exert influence to ensure that DG policies are followed. The DG committee should take on pain points that it can fix quickly with a noticeable improvement for the organization.

**SURVEY SAYS:**  
Executive mandate is crucial in early stages.

In an Internet-based survey TDWI Research conducted in 2007, 68% of survey respondents pointed to data ownership and other territorial issues as the leading barriers to data governance success. Overcoming these barriers requires a strong and attentive executive sponsor.

If the DG committee gets a couple of early successes, the rest of the child stage is a time of growth, as the committee expands its membership, develops more policies and procedures, and extends its influence to control data issues in more initiatives and applications.

A DG program grows in the child stage and gets more formal in the teenager stage.

The successful growth of the late child phase continues into the teenager phase, but in a better organized manner. For instance, the committee polishes its documentation for data usage policies and provides online mechanisms for people to get the documentation. It establishes procedures through which people can propose new policies and emendations to existing ones. Since data governance forces people to change how they use tools, applications, and data—and sometimes forces changes to the data itself—now is a good time to deploy procedures for proposing, tracking, and policing changes. The change management procedures may include workflow processes for review and authorization.

By the end of the teenager lifecycle stage, a successful DG committee is well-organized and has a broad reach into the data usage rules and standards for multiple business initiatives and technical implementations. But now the DG committee faces its greatest challenge—crossing the chasm to the adult stage. This will require the DG committee to embrace the Big Picture of enterprise-scope governance and meld with other governance bodies. Furthermore, changes to governance structure at the chasm must be coordinated with the requirements of other imperatives, especially the need for greater DG automation (as we'll see later in the discussion of imperative 4).

Sustaining growth in later maturity stages demands embracing the whole enterprise and its goals.

The adult lifecycle stage—just the other side of the chasm—tends to be about retooling for enterprise-scope governance. The retooling may force the DG committee to revise its membership to assure a broad enterprise representation. The DG committee may seek a new sponsor who has enterprisewide influence and may also revise policies and procedures to be more broadly applicable. If DG hasn't already been linked to other forms of governance, linkages occur or deepen in the adult stage in preparation for the federation of various governance types coming in the sage stage.

Few DG programs have reached the sage lifecycle stage of the TDWI Maturity Model, so it's difficult to say what's to be expected there. But a likely maturation for a successful DG committee is its governing the data usage practices of a long list of business initiatives and technical implementations. Another desirable maturation is DG becoming a subset of a larger governance program for IT governance, which in turn may be a subset of corporate governance. In fact, TDWI Research has encountered organizations that took a top-down approach to governance that caused

them to set up a centralized hierarchical relationship like this early on. When a DG program follows a bottom-up path from a single initiative (as most do), it takes three or more years for DG to become fully symbiotic with other forms of governance. Although the mature stages of DG focus on enterprise scope with centralized control, DG may still require regional variants, especially in multinational corporations.

## DG Imperative 2: Align with Data-Intense Business Initiatives

DG must support data goals of management-driven initiatives.

The second data governance imperative—*align with data-intense business initiatives*—concerns how a DG supports management-driven programs that rely on data as a critical success factor. Put more proactively, DG should guarantee the success of business initiatives by providing process and control for the data that the initiative uses or alters. When a business initiative requires data from multiple business units to be integrated, consolidated, or improved, the cross-functional team and change management processes of DG can be critical success factors. (See Figure 4.)

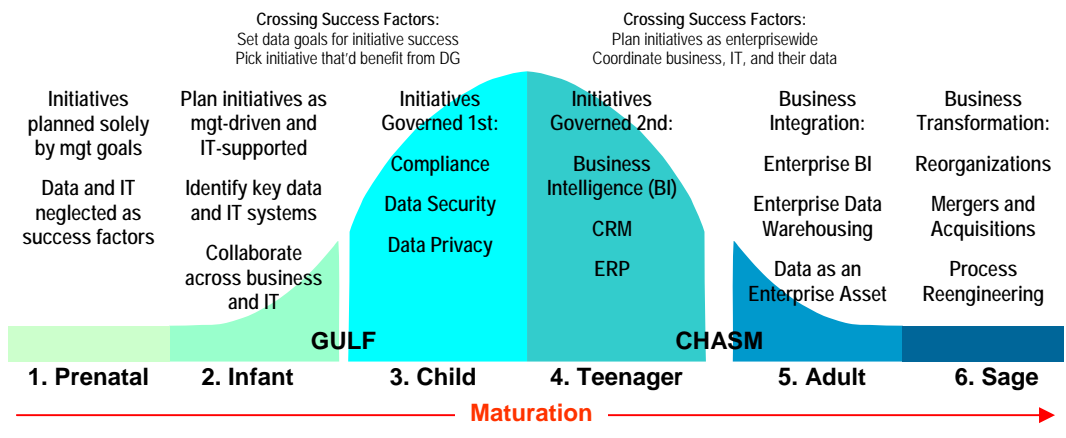


Figure 4. Maturity Model for DG Imperative 2: Align with Data-Intense Business Initiatives

Failing to govern IT systems and data can threaten business initiative success.

Most of us take for granted IT’s power to automate business processes, and we assume that data is one of the fuels that power the modern corporation. Unfortunately, there are still occasions where business management plans an initiative based mostly or solely on management goals, without due consideration of the role of IT and data. Many of the failed compliance initiatives publicized early this decade are good examples. And corporations continue to have unanticipated problems with mergers and acquisitions because due diligence and planning efforts failed to take into consideration the consolidation and integration of IT systems and data.

DG influenced by the second imperative sees management goals first, data and IT support of these second.

Organizations that make these mistakes are typically mired in the prenatal lifecycle stage of data governance. Organizations that have moved on to the infant stage or later conceive of business initiatives as being driven by management goals, yet fully supported by related IT systems and data. Once IT and data are recognized success factors for business initiatives, there arises a need for effective collaboration between business and IT people. When a corporate culture reaches these realizations, it may turn to data governance as a support mechanism for initiatives, because DG can provide an effective collaborative structure for data-intense business initiatives.

A data governance committee can also give a business initiative expertise for setting data-oriented goals, as well as processes and policies for effecting the changes that will achieve the goals. Once the business starts relying on DG this way, it’s halfway across the gulf to the child phase. To complete the crossing, the business needs to carefully select initiatives that will result in early successes by demonstrating DG’s ability to support management goals.

**Business initiatives for compliance, security and privacy are common starting points for DG.**

In the child lifecycle stage (where the first significant applications of DG occur), DG is regularly linked with compliance, data security, and data privacy, because DG can yield early and immediate successes. Compliance is often about (among other things) following regulations for data access and usage, and DG is built for this. Data security and data privacy concern establishing and complying with rules for who can access which data in which contexts; obviously, DG committees are designed for creating and enforcing data-oriented policies. Compliance and security/privacy initiatives may overlap (as they do with HIPAA compliance), and the DG committee can coordinate initiatives that might otherwise generate conflicting policies.

As DG proliferation continues from child to teenager stage, DG may extend its reach into business initiatives that are especially data-driven, like business intelligence (BI), customer relationship management (CRM), and enterprise resource planning (ERP). Although you may think of these as application types, the applications are ideally the outcome of business initiatives that seek to automate business processes for decision making, customer service and leverage, and operations. The business processes being developed handle a lot of data—much of it shared across organizational boundaries—and DG can provide policies and other guidance for this situation.

**SURVEY SAYS:  
88% surveyed say  
BI initiatives would  
benefit from DG.**

In a TDWI Best Practices survey of Q4 2007, respondents selected business intelligence far more often than other business initiatives as a prime target for data governance (88% of survey respondents). This makes perfect sense, because many business users consume corporate data that is delivered to them via reports and other media generated by BI systems, and DG can provide policies for this kind of information delivery. Furthermore, DG can deepen the IT-to-business collaboration that is a hallmark of BI programs.

Crossing the chasm successfully to the adult stage is, again, a matter of adjusting to enterprisewide governance, which—pertinent to the second DG imperative—means keeping DG and business initiatives properly aligned through a deeper business-to-IT collaboration. Assuming a successful crossing, in the adult stage DG can support so-called “business integration” initiatives, which are global programs that seek to unify the larger organization through changes to process and IT usage. Many business integration initiatives are extensions of BI and data warehouse programs, but taken to an enterprise extreme.

From a DG viewpoint, the most demanding business integration initiative is the “data as an enterprise asset” program. This can be highly disruptive when it takes data and applications away from their original owners and hands them to new owners or a central IT organization. The point is to pool data and increase access to it for the sake of broader enterprise visibility. Surviving this level of disruption and realignment is unlikely without DG and other forms of governance.

**In mature stages, DG supports recurring business integration and transformation initiatives.**

In the sage stage, the second DG imperative may lead to so-called “business transformation” initiatives, which invoke sweeping changes to organizational structures, as seen in reorganizations, mergers and acquisitions, and process reengineering projects. The point is to enable the enterprise to adjust quickly to changes in the business environment. Here, multiple forms of governance and other resources are brought to bear on fast-paced, recurring transformations. A mature organization will remember lessons learned in earlier lifecycle stages and leverage DG to plan business goals with full support from related IT systems and data. Furthermore, business integrations and transformations tend to force application and database migrations and consolidations in support of organizational changes, and DG can provide data standards and information lifecycle management (ILM) policies to guide database migrations and consolidations.



## DG Imperative 3: Govern Data Usage via Technical Implementations

Data management implementations and operational applications must respect DG policies.

When executed broadly, DG influences almost all data management practices, including data quality, integration, warehousing, standards, administration, architecture, and lifecycle management. Likewise, DG affects operational applications, because end-users access and use data through these. The third imperative for data governance—*govern data usage via technical implementations*—is about the relationship between data governance and specific applications and data management implementations. (See Figure 5.) DG typically requires that adjustments be made in these practices in support of the policies developed by the DG committee. Think of these practices as guidance on the specification level, not the design level. With end-user applications, the guidance may focus on how users create or alter data via applications’ user interfaces.

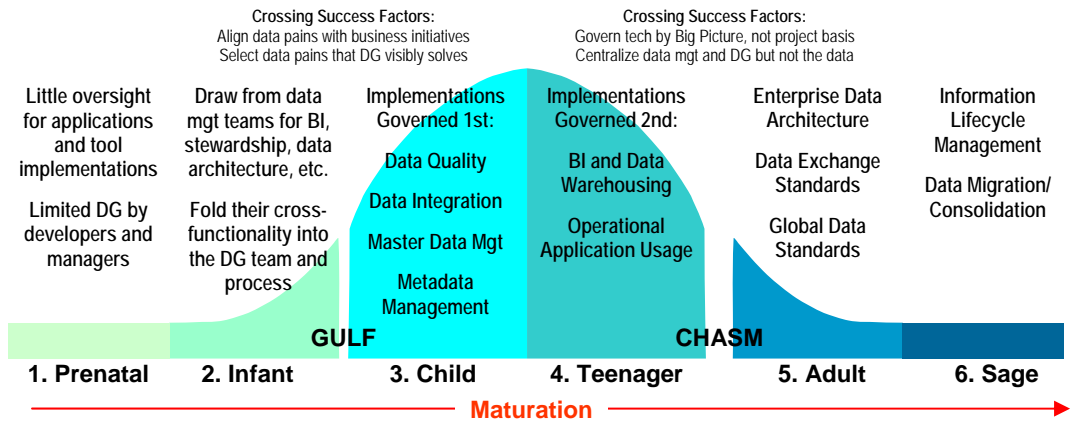


Figure 5. Maturity Model for DG Imperative 3: Govern Data Usage via Technical Implementations

As with other DG imperatives, the third one sees little true data governance (or oversight of any kind) relative to technical implementations in the prenatal lifecycle stage. It’s in the infant stage that diverse people begin to organize to solve specific data problems. Note that some data governance programs are pushed down from upper management (as we saw with the second imperative), while others rise up from grass-roots activities. The latter is certainly true when the third imperative wields greater influence than the second one. In these cases, technical people band together to create policies, procedures, and collaborative mechanisms for initiating or improving data management practices. Because of their technical background, these people draw from their experiences in stewardship, BI, and data architecture to form a cross-functional team.

DG influenced by the third imperative sees data pain first, business benefit second.

When driven by the third imperative, the data governance team may at first be cross-functional in the sense of incorporating people from multiple data management disciplines. Ultimately, though, it should also be cross-functional by including business people, who help assure business alignment. The fully staffed DG committee resulting from the third imperative tends to identify data problems first, then prioritize problems by their business pain. (This is backwards from how a DG committee driven by the second imperative begins with management goals and then depends on DG to support those goals.) For a technology-oriented DG committee like this, crossing the gulf into the child stage depends on business committee members guiding data management work into data problems that will yield recognizable business improvements when solved.

Once on the other side of the gulf, a different success factor kicks in. Getting various data management teams to do the work recommended by a DG committee is trivial in these situations, since members of these teams are usually members of the committee, as well. The problem is that

the technology-oriented DG team may dream up new data management solutions, which require new funding and staffing before the recommended work can begin.

DG may be founded to initiate new data mgt implementations, typically for DQ or MDM.

For example, TDWI Research has encountered numerous organizations that founded a DG committee to assure the success of data quality (DQ) or master data management (MDM) implementations. This makes sense, because both DQ and MDM are strongly cross functional. You need technology people to identify specific data pain points and business people to prioritize them. Also, DQ and MDM are cross functional in that improvements to the quality of data or master data definitions almost always span across multiple IT systems and the departments that own them. DG excels in coordinating and coercing people involved in data changes and improvements—especially when these span multiple departments and/or IT systems—which explains why DG is so often associated with DQ and MDM technical implementations. However, when DG predates these implementations, the DG committee needs a business sponsor or influential members who can provide new resources for these new solutions.

**SURVEY SAYS:**  
Data integration and quality implementations need DG guidance.

A TDWI Technology Survey of August 2007 asked: “Which data management practices do you think should be guided by data governance?” At the top of their picks, survey respondents chose data integration and data quality (83% and 82%, respectively). Data integration (DI) implementations are ripe for DG controls, because they “use” data by accessing and transporting it, often across organizational boundaries. DQ has become almost synonymous with data governance because it’s a critical success factor in DG-driven initiatives for compliance, business transformation, and business integration.

After DI and DQ, survey respondents ranked MDM and metadata management as also benefiting from DG. Other technical implementations addressed by DG relatively early in its maturation include BI platforms, data warehouses, and various operational applications and their data.

Crossing the chasm always entails an embrace of enterprise scope. With the technology oriented third DG imperative, success factors relate to the Big Picture and centralization. For example, in the adult stage, DG typically coordinates with enterprise data architecture personnel to improve the Big Picture of enterprise data by developing a variety of standards and policies for data exchange, data models, and data interfaces. In the sage stage, this development may mature into global policies for the ILM of databases and their platforms (not just data archive procedures, which most ILM is stuck on today). This mature, governed approach to ILM ably assists with the life-and-death decisions made about applications and databases that occur in concert with the business integrations and business transformations described in the discussion of the second DG imperative.

DG Imperative 4:

## Automate DG Processes via Technical Implementations

Some data mgt tools can automate some actions of DG.

As we’ve just seen, DG is inherently organizational and interpersonal, even when it governs data usage via technical implementations. However, many of the tasks of the governance process—and many of the outcomes that result from enforcing a governance policy—can be automated (to some degree) with computer software and hardware. Automation is important, because it helps a maturing DG program communicate its policies to affected parties, collaborate with DG committee members, enforce its data usage policies, and grow into more initiatives and implementations governed. Different forms of automation kick in at different lifecycle stages, so the maturation of DG automation is what the fourth DG imperative is all about.

**SURVEY SAYS:**  
 More than half of users think software can automate DG processes.

Is software automation for DG really possible, though? More than half of the respondents to a TDWI Best Practices survey said yes, a quarter said no, and the rest don't know, indicating that some kind of software automation for DG is possible. TDWI's opinion is that some data governance tasks can be automated with software, and some can't, as explained below.

Common goals for any data-oriented software automation include the consistent, scalable, and auditable management, repurposing, and communication of information. With that in mind, there is noticeable overlap between the goals of DG and the capabilities of various data management tools. In particular, most tools for data integration and quality have useful functions, as do some databases, metadata repositories, and operational applications (for ERP, CRM, and so on). The overlap suggests that such tools can help automate DG processes and outcomes.

At the moment, however, software automation specifically designed for DG is somewhat light. A few selected functions or areas within certain types of data management and development tools support DG, but full-blown applications specifically for data governance are rare. TDWI Research suspects that the available automation for DG will increase over the next year or two, because many users are in the midst of defining their requirements for DG software automation, and vendors are already expanding data management products to address DG more directly.

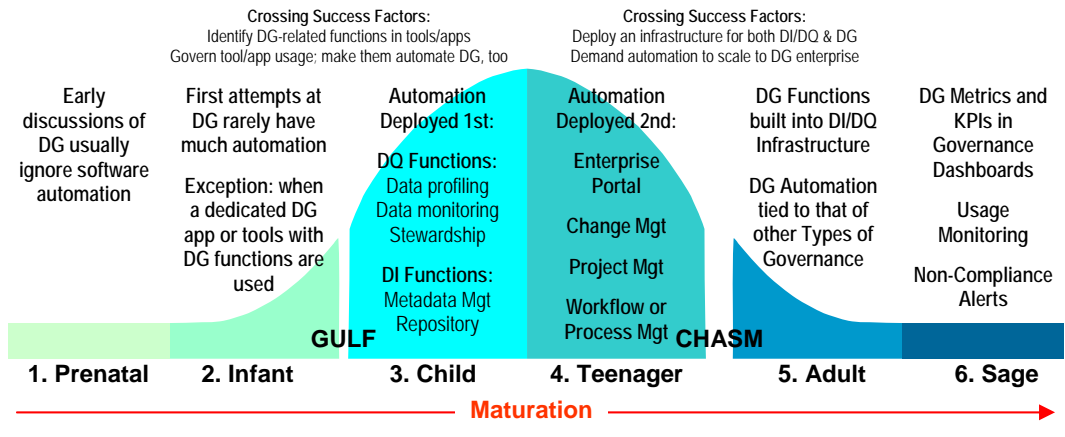


Figure 6. Maturity Model for DG Imperative 4: Automate DG Processes via Technical Implementations

Turning now to the TDWI Maturity Model, the fourth DG imperative charts the maturation of DG automation via software (see Figure 6). In the prenatal lifecycle stage, early discussions of data governance are so preoccupied with defining the DG process that they ignore its automation via technical implementations or otherwise. Even in the infant stage—where an organization first attempts proof-of-concept projects—DG is still focused on organizational issues (like populating the committee and drafting policies), so software automation is minimal, if present at all.

As an exception, some organizations decide early on to use a dedicated data governance application and hence have automation from the beginning. This may be the case in organizations that have a large packaged application for ERP or CRM, and they simply purchase an additional module or turn on extra functionality for DG. In other cases, the organization may acquire an application for compliance or risk that happens to have DG or other governance functions embedded. However, the majority of organizations today build up their software automation for data governance by progressively using more of the functions built into pre-existing tools and applications—whether those functions were designed for governance or not.

Ironically, a DG program relies on the IT systems it governs for much of its automation.

In fact, in the fourth DG imperative, one of the critical success factors for crossing the gulf is to identify DG-related functions in data management tools and operational applications that will apply directly to the DG solutions to be rolled out in the child phase. This reveals a curious irony inherent to data governance: a successful DG program will both govern data-related usage of tools and applications *and* press those same tools and applications into service automating DG tasks.

In the child lifecycle stage, DQ tools play a crucial role in DG automation. DQ tools support a variety of quality operations for name-and-address cleansing, match-and-merge, deduplication, verification, enhancement, standardization, and so on. For DG programs that are oriented to data improvements or quality methodologies in general (such as Six Sigma and Zero Defect), DQ tools are indispensable for achieving quality goals for enterprise data. But DQ tools also include capabilities that have direct import to the data governance process.

DG automation can tap DQ tool functions for profiling, monitoring, and stewardship.

For example, *data profiling* helps a user discover data and quantify its state; this area within the tool originated for data quality purposes but can be applied to inventorying data assets and assessing their need for data governance. Likewise, *data monitoring* polls data after each run of a deployed data quality solution to assess whether the data is improving and to identify further opportunities for improvement; monitoring can assure compliance with DG policies, not just data quality standards. And the practice of *data stewardship* is so prominent in data quality implementations that most data quality tool vendors have created tools (or functions within a data quality tool) designed specifically for the steward as user. These tools enable the steward to discover data improvement opportunities, develop rules for data transformations, communicate these to technical developers, review reports based on data monitoring, and process exceptions with a mix of manual and automated methods. Again, all these stewardship capabilities have direct application to DG automation.<sup>2</sup>

Similar to DQ tools, DI tools make a significant contribution to the automation of DG. DI tools support various data access, movement, and transformation functions, which are applied to many of the technical implementations governed by a DG committee. Less obvious, however, are the semantic data facilities built into DI tools that can assist with DG.

DG automation can tap DI tool functions for metadata management.

For example, many tool types are capable of metadata management, to a certain degree. But most DI tools have an advanced facility for metadata management, which includes metadata extraction and integration features, plus a metadata repository that can manage metadata, documentation, development artifacts, collaborative documents, project management timelines, data lineage information, and so on. DI tools must interoperate with many types of tools and applications, and most of the interoperability depends on shared metadata. So, the metadata facility in most DI tools is open, such that it can manage metadata and other objects for its own DI tool, as well as for many other tools and applications. Because of their advanced features and openness, the metadata management and repository functions of DI tools are often tapped to be the central repository for projects in BI, data warehousing, data quality, ETL, master data management, and other data management implementations. Now, users are tapping this facility to automate DG.

DG automation requires advanced management for semantic data.

In particular, metadata management functions help build an inventory of governed data and define the meaning of governed data. The semantics may be expanded to define other attributes, like data's owners, sources, transformations, targets, quality state, dependencies, security issues, and so on. The data inventory and definitions are crucial to DG goals such as business integration, data quality, and auditability. Given the broad capabilities of DI tools' metadata repositories, they can

---

<sup>2</sup> For a complete survey of data quality vendors and tools, see the TDWI Technology Market Report *Enterprise Data Quality Tools* (Q2 2006), available online at [www.tdwi.org/research](http://www.tdwi.org/research).

also automate DG collaboration by managing policy and procedure documents. When DG goes so far as to audit technical implementations, DG auditors may review development artifacts (such as jobs and routines) managed by the centralized repository.

Although an organization begins its DG automation by pressing DQ and DI tool functions into service (typically in the child lifecycle stage), it soon thereafter turns to automation for DG collaboration and workflow (in the teenager stage or later).

DG processes generate documents that need management, perhaps in a workflow.

Like any corporate program, data governance generates a lot of paperwork through which people collaborate. This includes policy documents, forms for change requests and other procedures, board descriptions, documents chronicling board decisions, project timelines, meeting minutes, and a plague of other documents. It's best to store these in a central place where everyone can access the latest document version as they need it. The mechanism for sharing Microsoft Office files (which most of them are) can be as simple as a folder on a shared network drive or as complex as a full-blown enterprise portal. The work of the DG committee involves the review, authorization, and audit of many change proposals. So, when great complexity or volume is involved, it makes sense to automate DG's change management processes with applications for project management, workflow, or process management.

As with other imperatives, crossing the chasm to the adult stage of maturity involves taking data governance to the enterprise. For the automation-oriented fourth DG imperative, this means much more automation for DG processes, just to survive the growth and scalability requirements of the enterprise. Hence, adoption trends begun in the child and teenager stages ramp up considerably in the adult and sage stages, as even more DQ and DI tool functions are applied to DG.

In a related trend, leading vendors are well on their way to producing unified platforms that support DI, plus a host of related data management practices, like DQ, data profiling/monitoring, stewardship, metadata management, master data management, ETL, data federation, replication, and so on. By 2010 or so, these unified platforms will constitute a DI/DQ infrastructure that most enterprise data will travel. So users should plan to tap the upcoming DI/DQ infrastructure for even more functions that are conducive to data governance.<sup>3</sup>

Advanced automation may present DG metrics or KPIs in dashboards.

In the adult and sage stages, DG committees and processes must meld with those for corporate and IT governance, and so DG automation must, too. Users whom TDWI Research has interviewed have described how they would like to apply performance management methods to governance, once it reaches maturity. Hence, a possible outcome concerning automation is that metrics for DG and other forms of governance may be monitored and managed from dashboards. Behind the dashboards, various usage monitoring tools could collect statistics about data and its use, and this data would populate dashboard metrics. Data, tools, applications, or departments found to be noncompliant with governance policies could be highlighted in the dashboard, or an alert could be sent to relevant parties. Since multiple forms of governance relate in a hierarchy, a hierarchical approach to performance management could be appropriate, as with scorecards and cascading dashboards. Of course, few organizations have reached a fully mature level of automation for data governance (or any form of governance), so it's difficult to say whether this vision of advanced DG automation will come true.

---

<sup>3</sup> For a complete survey of data integration vendors and tools (with a focus on unified DI platforms), see the TDWI Technology Market Report *Data Integration Tools* (Q4 2007), available to TDWI Members at [www.tdwi.org/research](http://www.tdwi.org/research).

# The TDWI Maturity Model for Data Governance

The four DG imperatives interact in a single DG Maturity Model.

We’ve seen that a data governance program unfolds over time through a complex series of events that involve the coordination of many parties and technologies. Yet the complexity is comprehensible because DG’s components organize into four imperatives that mature through six lifecycle stages and two gaps. The TDWI Maturity Model provides structure for the lifecycle stages, although data governance doesn’t compress conveniently into a single maturation thread.

The chief insight of this TDWI Monograph is that data governance is a multithreaded process. Each DG imperative is a distinct thread that unfolds through (1) organizational issues, (2) business initiative alignment, (3) technology implementation governance, and (4) automation with software, respectively. All imperatives progress through the same lifecycle stages, and—despite variations—the threads weave together in the cloth of a single maturity model for DG. (See Figure 7, which combines the imperative-oriented maturity models of Figures 3–6.)

|  | 1. Prenatal   | 2. Infant  | GULF  | 3. Child   | 4. Teenager   | CHASM   | 5. Adult  | 6. Sage  |
|--|---|--|---|--|---|---|---|--|
| <b>Imperative No.1:</b><br>Maintain a Cross-Functional Team and Process        | Ad Hoc Governance<br>Steering Committees<br>Stewardship<br>Per-System Data Policies                 | DG Team Version 1.0<br>Governance Prototyping<br>Rally around Beachhead  | Get executive mandate.<br>Focus on small-scope fixable pains.                                     | Define Policies for Data Usage<br>Get Early Successes<br>Revise DG Team<br>Govern More Stuff   | Improve Processes<br>Change Mgt Authorization<br>Workflow<br>Govern More Stuff                        | Embrace enterprise Big Picture.<br>Meld with other governance teams.                          | Revise Team, Procedures, Policies, for Enterprise<br>Link DG to other forms of Governance               | Centralized Hierarchical Governance<br>Regional Variants<br>Seek Global Architecture           |
| <b>Imperative No.2:</b><br>Align with Data-Intense Business Initiatives        | Initiatives planned solely by mgt goals<br>Data and IT neglected as success factors                 | Plan initiatives as mgt-driven and IT-supported<br>Identify key data and IT systems<br>Collaborate across business and IT            | Set data goals for initiative success.<br>Pick initiative that'd benefit from DG.                 | Initiatives Governed 1st:<br>Compliance<br>Data Security<br>Data Privacy   | Initiatives Governed 2nd:<br>Business Intelligence (BI)<br>CRM<br>ERP                                 | Plan initiatives as enterprise wide.<br>Coordinate business, IT, and their data.              | Business Integration:<br>Enterprise BI<br>Enterprise Data Warehousing<br>Data as an Enterprise Asset    | Business Transformation:<br>Reorganizations<br>Mergers & Acquisitions<br>Process Reengineering |
| <b>Imperative No.3:</b><br>Govern Data Usage via Technical Implementations     | Little oversight for applications and tool implementations<br>Limited DG by developers and managers | Draw from data mgt teams for BI, stewardship, data architecture, etc.<br>Fold their cross-functionality into the DG team and process | Align data pains with business initiatives.<br>Select data pains that DG visibly solves.          | Implementations Governed 1st:<br>Data Quality<br>Data Integration<br>Master Data Mgt<br>Metadata Management  | Implementations Governed 2nd:<br>BI and Data Warehousing<br>Operational Application Usage             | Govern tech by Big Picture, not project basis.<br>Centralize data mgt & DG, but not the data. | Enterprise Data Architecture<br>Data Exchange Standards<br>Global Data Standards                        | Information Lifecycle Management<br>Data Migration/Consolidation                               |
| <b>Imperative No.4:</b><br>Automate DG Processes via Technical Implementations | Early discussions of DG usually ignore software automation  | First attempts at DG rarely have much automation<br>Exception: when a dedicated DG app or tools with DG functions are used           | Identify DG-related functions in tools/apps.<br>Govern tool/app usage; make them automate DG, too | Automation Deployed 1st:<br>DQ Functions:<br>- Data profiling<br>- Data monitoring<br>- Stewardship<br>DI Functions:<br>- Metadata Mgt<br>- Repository | Automation Deployed 2nd:<br>Enterprise Portal<br>Change Mgt<br>Project Mgt<br>Workflow or Process Mgt | Deploy an infrastructure for both DI/DQ & DG.<br>Demand automation to scale DG to enterprise. | DG Functions built into DI/DQ Infrastructure<br>DG Automation tied to that of other Types of Governance | DG Metrics and KPIs in Governance Dashboards<br>Usage Monitoring<br>Non-Compliance Alerts      |

Figure 7. The Four DG Imperatives combined in a single DG Maturity Model

If the multithreading metaphor doesn’t appeal to you, you could think of data governance as a multidimensional modeling problem, as are most business entities and programs. The lifecycle stages and gaps of TDWI’s maturity model define the time dimension, whereas the imperatives drill into dimensions representing business and technical domains. Regardless of the metaphor you prefer, the point is that a data governance program of any size or maturity consists of multiple imperatives that have individual requirements and need coordination.

## Recommendations

DG is multidimensional by nature, so plan and execute it accordingly.

The multithreaded nature of the larger DG process has ramifications for real-world DG programs:

- **Plan the DG program as four imperatives that interact.** This approach reduces the complicated program into manageable chunks to avoid a “big bang” approach. If including four imperatives is too fine-grained for your organization’s planning style, you might group the imperatives into two pairs (as in Figure 1) for organization and technical issues, respectively.
- **Coordinate the four DG imperatives to achieve a unified DG program.** This may be as simple as revisiting the alignment of imperatives when the overall DG program approaches key milestones, like the lifecycle stages and gaps. Since the DG committee is a cross-functional mix of people, committee members could be assigned to specific imperatives and could coordinate their plans and policies as part of the committee’s cross-functional collaboration.

Coordinating the four imperatives of DG is fairly straightforward because all imperatives share a few common goals and success factors on a per-lifecycle-stage basis:

Plan your DG program with the TDWI Maturity Model, and execute its details via the DG imperatives.

- **Get organized and prove the DG concept in the prenatal and infant stages.** Organizing the DG committee (or its equivalent) must take priority here, followed by proto-governance activities (especially data stewardship) that prove the value of governance for data, whether with business initiatives or technical implementations.
- **Carefully choose a leader and first targets before attempting to cross the gulf.** You’ll need an influential executive sponsor to give DG clout, and you’ll need people’s buy-in for the list of initiatives and implementations to which you’ll first apply DG.
- **Stick to agreed early targets, but apply DG flexibly in the child and teenager stages.** Expect to revise team membership and every policy and procedure. If you find the right mix for your corporate culture, DG will proliferate due to greater relevance and repeatability. Don’t forget to ramp up DG automation, not just the organizational DG process.
- **Prepare to embrace the enterprise before attempting to cross the chasm.** By this point, you’ll have some automation in place; enterprise volume will require much more. Likewise, you should already be coordinating DG with other governance bodies; expect to devote a lot of resources and bandwidth to more coordination as governance unifies across the enterprise.
- **Execute on enterprise DG goals, supported by enterprise infrastructure.** Don’t expect DG to deliver global goals by itself, though. DG in adult and sage stages should ably assist enterprise-oriented initiatives (like business integrations and transformations) and implementations (like enterprise data architecture or data as an enterprise asset). Enterprise success requires a great deal of DG automation via a comprehensive infrastructure for data integration and quality.

Finally, be creative as you apply the TDWI Maturity Model for data governance:

Use the TDWI Maturity Model for DG as inspiration and a framework.

- **Tailor the model to your organization.** After all, each organization has unique requirements due to its mix of culture, technology, industry, region, and so on. The model shown here is intentionally symmetric to facilitate readers’ comprehension, with milestones are plotted roughly. Your adaptation of the TDWI Maturity Model will no doubt be asymmetric, with a different arrangement of milestones and possibly a different number of lifecycle stages or

imperatives. Even so, this Monograph's interpretation of the TDWI Maturity Model for data governance should give you many ideas and a framework for the short-term conception and long-term planning of your organization's data governance program.