

Metadata Management for Holistic Data Governance

Using Informatica Metadata Manager & Business Glossary
to Govern Your Data and Deliver Better Business Decisions

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 July 2013

Table of Contents

Table of Contents	1
Executive Summary	2
The Role of Metadata and Its Management in Data Governance	3
Data Governance: The Means, Not the End	3
What Is Metadata?	3
The Importance of Effectively Governed Metadata	4
The Importance of Metadata Management in Data Governance	5
Business Drivers for Metadata Management in Data Governance	6
Special Consideration: Big Data and Data Governance	7
Informatica Metadata Manager & Business Glossary for Holistic Data Governance	7
Informatica Metadata Manager & Business Glossary Features	7
The Metadata Manager Interface	9
The Business Glossary Interface	11
Practical Considerations for Metadata Management	13
Best Practices	13
Sizing Your Metadata Effort for Data Governance	15
Other Important Tools for Data Governance	16
Conclusion	17

Executive Summary

The ultimate goal of any data governance initiative is to manage data for the purpose of delivering trustworthy, timely, and relevant information in support of informed business decisions.

Data governance, then, is not about data, per se: “it is about the business processes, decisions, and stakeholder interactions that you want to enable.”¹ Once an enterprise establishes the making of informed business decisions as the collective aim of data governance, it becomes much easier to define clear goals and priorities for an initiative. Good data governance begins with specifying the company policies, people, and processes needed to launch that initiative. The focus then turns to the tools and technologies to implement and ensure compliance with data governance requirements.

Key to successful data governance is the management of metadata—the frame of reference giving data its context and meaning. Effectively governed metadata provides a view into the flow of data, the ability to perform an impact analysis, a common business vocabulary and accountability for its terms and definitions, and finally an audit trail for compliance. The management of metadata becomes an important capability enabling IT to oversee changes while delivering trusted, secure data in a complex data integration environment. Good metadata management tools, then, play a central role in holistic data governance.

This white paper makes the case for metadata management as a core foundation for any data governance initiative. It discusses how Informatica® Metadata Manager & Business Glossary can contribute significantly to holistic data governance by providing a high degree of control over metadata along with the flexibility to deal with new and changing requirements. Informatica Metadata Manager & Business Glossary is a single product with a shared metadata repository and specific interfaces to meet the different needs of business and technical users:

- **Metadata Manager** enables IT staff to manage technical metadata
- **Business Glossary** enables business and IT stewards to collaboratively manage business metadata

The paper also explains features of Metadata Manager & Business Glossary that significantly boost the productivity of IT and business staff collaborating on projects. Collectively, these features contribute directly to enterprise agility and are a major influence on the ultimate success of data governance efforts.

¹ Rob Karel, VP Product Strategy, Informatica Corporation

The Role of Metadata and Its Management in Data Governance

Data Governance: The Means, Not the End

At Informatica, we define data governance as “the functional coordination and definition of processes, policies, standards, technologies, and people across the organization to manage data as a corporate asset. This enables managing the availability and controlled growth of accurate, consistent, secure, and timely data for better decision making, reduced risk, and improved business processes.”

Data governance focuses on the delivery of trustworthy, secure information to support informed business decisions, efficient business processes, and optimal stakeholder interactions. It is therefore not an end in itself, but merely the means: data governance supports your most critical business objectives.

A data governance program begins with a definition of its strategy and goals as well as the scope of its success. Next, there is the need to define the three data governance pillars, which are required to actually implement the initiative:

- Policies
- People (and people skills)
- Processes

Only after these initial steps are carried out is an organization in a position to determine the best tools and technology to implement its data governance initiative—including those necessary to manage metadata, once its role has been established.

What Is Metadata?

Metadata is what gives your data a frame of reference. Forrester Research defines metadata as “the information that describes or provides context for data, content, business processes, services, business rules and policies that support an organization’s information systems.”

For example, Apple Corporation’s App Store sells software applications online. The data in this case is the applications, or apps. The metadata is the information about those apps—an app’s description, price, user rating, reviews, and developing company.

In the context of data management environments, there are several relevant types of metadata:

- **Technical metadata** provides technical information about the data, such as the name of the source table, the source table column name, and the data type (e.g., string, integer)
- **Business metadata** supplies the business context around data, such as the business term’s name, definition, owners or stewards, and associated reference data
- **Operational metadata** furnishes information about the use of the data, such as date last updated, number of times accessed, or date last accessed

The Importance of Effectively Governed Metadata

Business, technical, and operational metadata provide the context and meaning to turn data into information. Put simply, Data + Context = Information.

A customer in the health insurance industry recently shared that it was experiencing data quality issues. The problem was not that the data itself was bad, but that the meaning of the data varied. In this case, the term “Claim Paid Date” could mean any of the following:

- Date the claim was approved
- Date the check was cut for the claim
- Date the check cleared

All of these were valid definitions of “Claim Paid Date,” yet their use cases were very different. This organization needed to define when to use which definition and to give each use case a slightly different, unique business term to avoid confusion. Its solution focused on building an enterprise business glossary to define its core terms and to kick off a data governance initiative. This effort clarified the meaning of terms and provided standards defining the business context in which to use each term.

In addition to resolving this type of data ambiguity, effectively governed metadata provides:

- **Visibility into how data is flowing through the environment.** The first step toward managing the environment is to have a clear visual map of how the data is moving between data sources and targets.
- **An impact analysis and root cause analysis.** Impact analysis enables IT staff to see the impact of a proposed change to the environment before it is implemented and to understand the complex cross-dependencies involved. (For example, changing the definition of “Due Date” requires the ability to see what other applications, reports, and analyses will be impacted by the change.) Root cause analysis enables business users to drill back from a term in a report to understand the source of the data and how it was moved and transformed as it was added to the report.
- **A common business vocabulary that standardizes terminology.** The resulting business glossary enables clear communication among business units and between business and IT.
- **Accountability into who is responsible for business terms and definitions.** One of the core concepts of data governance is accountability. If an issue is discovered relating to a business term, it is important to be able to identify who owns the term and is responsible for remediating any issues. It is also important to be able to see who has made changes to the term and when it was changed.
- **Audit trail for auditability and compliance.** Metadata management should be able to tell you who changed what and when. It should also be able to provide timely answers to questions about the flow of data in the environment. (For example, the audit trail should be used to easily determine from where a particular number originated and how it was calculated.) The transparency is especially important to ensure compliance with many data-driven regulatory edicts such as Sarbanes-Oxley and Basel II.

The Importance of Metadata Management in Data Governance

As one executive at a major bank has noted, “Any metadata management initiative without data governance is going to fail.” Metadata management can be an important capability enabling IT to manage change while delivering trusted, secure data in a complex data integration environment. The benefits become compelling when business stakeholders systematically engage in this process and accept ownership of the reference frame around data. At that point, enterprises can link the business metadata with the underlying technical metadata, providing a vocabulary and context for collaboration across the company.

For example, when a business user asks an IT counterpart to show “net revenue” in a report or analysis, it will no longer prompt the question, “which net revenue—financial, sales, or marketing?” Good metadata management contributes to data governance by helping to pre-empt such important questions, among several others:

- What does this business term mean?
- Which (of several similar) business terms should I use?
- What is the source(s) of the data behind that term?
- How was the data transformed as it was moved from source to target?
- Who is accountable for the definition, documentation, and management of that term?
- Who has changed this term? How and when was it changed?
- What policies and rules are applicable to that term? (Examples include data quality rules, security masking rules, archiving rules, and data retention policies)
- What other data objects would be impacted by a change to a specific data object in the environment?
- How long will it take to implement a change to the environment, without causing issues to other reports and analyses that may use the same data objects?

Business Drivers for Metadata Management in Data Governance

A number of corporate initiatives drive efforts in data governance—and with it, metadata management. These initiatives include:

- **Common business vocabulary** (simple data stewardship). This is a start-small approach that focuses on a common business vocabulary for a specific problem space or business unit.
- **Full data governance** (or data management strategy). A more top-down approach, this is typically a larger program involving a number of business units across the enterprise and targeting various business opportunities in a program managed over multiple phases, if not longer.
- **Regulatory compliance.** These initiatives are driven by the need to comply with international, national, local, or industry mandates. Regulatory compliance—usually managed within a governance, risk, and compliance (GRC) function—is obviously interdependent with data governance. While discovering, analyzing, and documenting many of the organization’s internal data governance requirements, it also must reconcile requirements specific to the external regulation being applied. Some examples of these include:
 - Banking: Basel II, Basel III, Dodd-Frank Act, money laundering
 - Insurance: Solvency II
 - Healthcare: HITECH Act, HIPAA
 - General Financial: Sarbanes-Oxley Act
- **Metadata Management.** This is a more incremental approach where metadata management and data governance are attached to each new business initiative as “best practices.” The initiative defines the business case and scope of the project. This approach has found success at a number of companies where larger data governance initiatives have failed. Benefits of this approach:
 - Owners of the business initiative own the scope and prioritization
 - Owners of the business initiative are responsible for the justification and budget
 - It reduces risk of scope-creep, which can make the effort less relevant

Examples of business initiatives could include projects such as a proactive customer engagement initiative involving MDM, CRM, and services capabilities, or a data security initiative that involves data access, data masking and privacy, and data retention requirements.

Special Consideration: Big Data and Data Governance

Virtually all enterprises are faced with the challenge of managing the volume, velocity, and variety of data. Hadoop/MapReduce technology brings some interesting advantages in terms of the ability to handle complex data analytics and to enable extreme data scalability at relatively low cost. It is unlikely that Hadoop will replace relational DBMSs any time in the near future; rather, the two technologies will likely co-exist, as each offers unique strengths.

While the technologies used to manage and analyze the data may vary, the goal of metadata management and data governance should remain the same: to deliver trustworthy, timely, and relevant information in support of good business decisions. There is no such thing as “Big Data Governance” or “Big Data Metadata Management”—instead, it’s a matter of extending holistic enterprise data governance and metadata management efforts to include new data types and sources.

One of the challenges that Hadoop presents is metadata management. Without good metadata management and data governance, Hadoop will lack transparency, auditability, and the ability to standardize and reuse data. Enterprises will still need visibility into critical information about the data, such as its origin, quality, and ownership, or risk having Hadoop become another island of data in their environment. New technologies are emerging in this space such as HCatalog and Hive/HiveQL, which will make it easier to harvest metadata from unstructured and semistructured data, enabling data lineage for Hadoop. These capabilities are essential to integrating Hadoop technologies into the overall data integration framework so big data does not become isolated in the enterprise and can be governed like any other data source.

Informatica Metadata Manager & Business Glossary for Holistic Data Governance

Informatica provides a versatile and robust tool with precisely these capabilities to deliver trustworthy, secure data and launch successful metadata management initiatives. Its Metadata Manager & Business Glossary offers unique benefits that enable IT managers to minimize the business risk of damaging critical business data while implementing changes.

Informatica Metadata Manager & Business Glossary is a key component of Informatica PowerCenter® Advanced Edition. It provides the core metadata management tools necessary to set the foundation for a data governance initiative. Metadata Manager & Business Glossary is a single product with a shared metadata repository. It has two user interfaces for two distinctly different audiences:

- **Metadata Manager** enables IT staff to handle the technical metadata.
- **Business Glossary** allows business and IT stewards to collaboratively manage business metadata.

Informatica Metadata Manager & Business Glossary Features

In addition to the functions just mentioned, Metadata Manager & Business Glossary includes the following features that significantly boost the productivity of IT and business staff collaborating on projects. Together, these features contribute directly to enterprise business agility and play a major role in the ultimate success of data governance efforts.

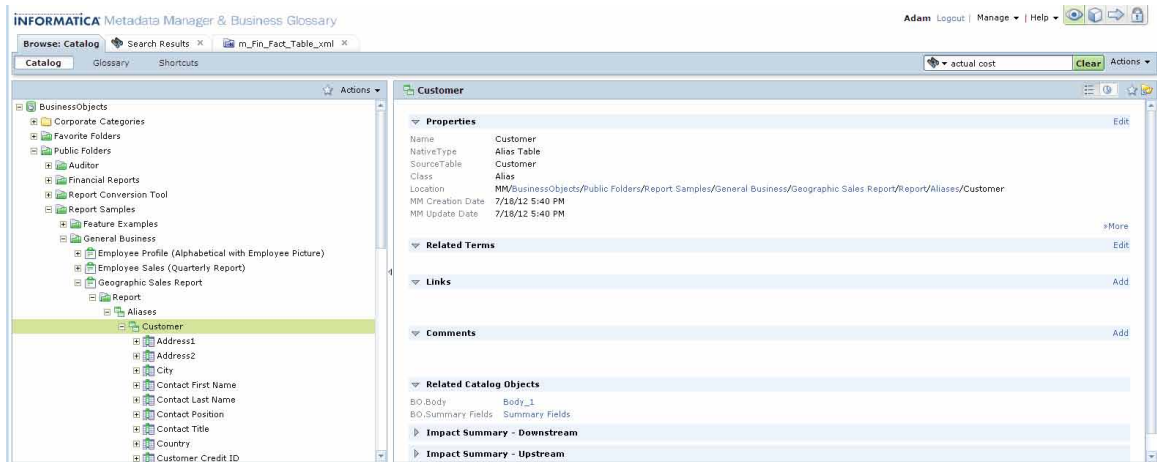


Figure 1: Metadata Manager provides a view of the technical metadata for IT professionals.

- **Search & Browse.** Informatica Metadata Manager & Business Glossary comes with a powerful search capability that includes advanced filtering. It also supports the ability to browse data sources in a folder/tree structure (see Figure 1). With this feature, staff can quickly find the data objects they are looking for.
- **Personalization.** Most data management environments become large and complex. Informatica Metadata Manager & Business Glossary supports the ability to filter out data that is not of interest from searches and views. For example: as a means of simplifying its environment, users of Cognos could choose not to view information about Business Objects or Microstrategy.
- **Linking of business and technical metadata.** Informatica supports the easy linking of business terms to the underlying technical metadata. This is where the “magic” occurs. Once the terms and technical metadata are linked, IT and the business staff finally share a common language for clear communications and collaboration between the organizations.
- **Collaboration.** Informatica Metadata Manager & Business Glossary includes built-in tools for annotation and messaging to optimize collaboration between business and IT. In addition, every view in Metadata Manager & Business Glossary offers a unique URL that can be shared in these communications, providing rich context to any conversation.

“Any sufficiently advanced technology is indistinguishable from magic.”

– Arthur C. Clarke, Author
2001: A Space Odyssey

The Metadata Manager Interface

Visibility

The Metadata Manager interface includes a **data lineage** function that provides a visual map of the flow of data through the data integration environment. Metadata Manager collects this metadata from enterprise applications, repositories, databases, flat files, PowerCenter, data modeling tools, and business intelligence tools. It visually renders the information, which can then be expanded to provide greater detail (see Figure 2).

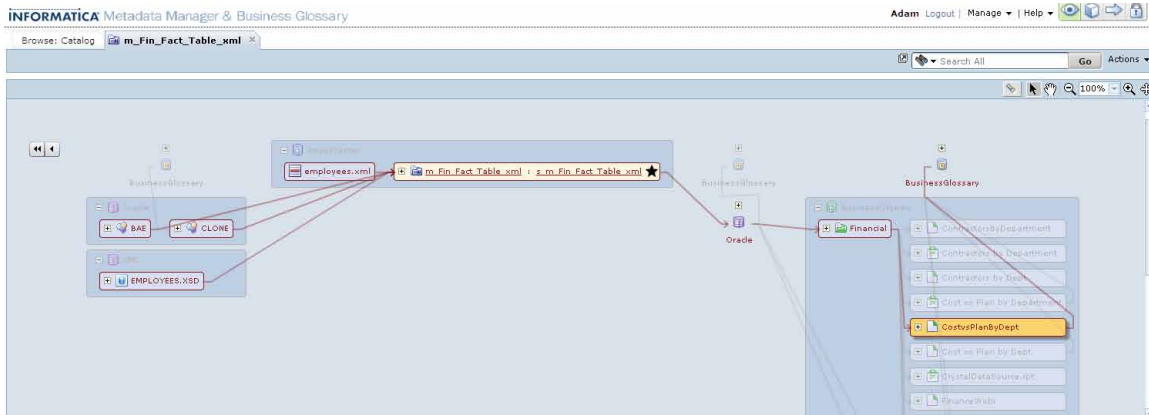


Figure 2: With data lineage, the Metadata Manager interface automatically creates an end-to-end view of the data flows through a data integration environment.

You can't manage what you can't measure—or in this case, what you can't visualize. So, data lineage becomes an important tool for data governance, providing visual documentation of how the current data integration environment is working. There are several specific uses for data lineage:

- Change management: for understanding the environment before a change is made
- Resolving data questions, such as, "This number in my report does not look right; can you confirm that is it correct?"
- Auditing and compliance: for allowing auditors to trace data from source to target in order to ensure that the data was calculated correctly and comes from an accurate, trustworthy source

As a whole, data lineage provides a view of the entire data integration environment. If you select a data object in that environment, the bold lines indicate the item's data lineage, from source to target. This is essential to understanding the data's source, how it was transformed as it moved through the environment, and where it is ultimately used (e.g., reports, dashboards).

Change Management

Once you deliver the overall visibility required across your data management environment, your next step should be to enable the analysis of the impact of a proposed change—before it is implemented. Metadata Manager includes the ability to perform **impact analyses** and root cause analyses to understand where the data came from and the impact of proposed changes (see Figure 3).

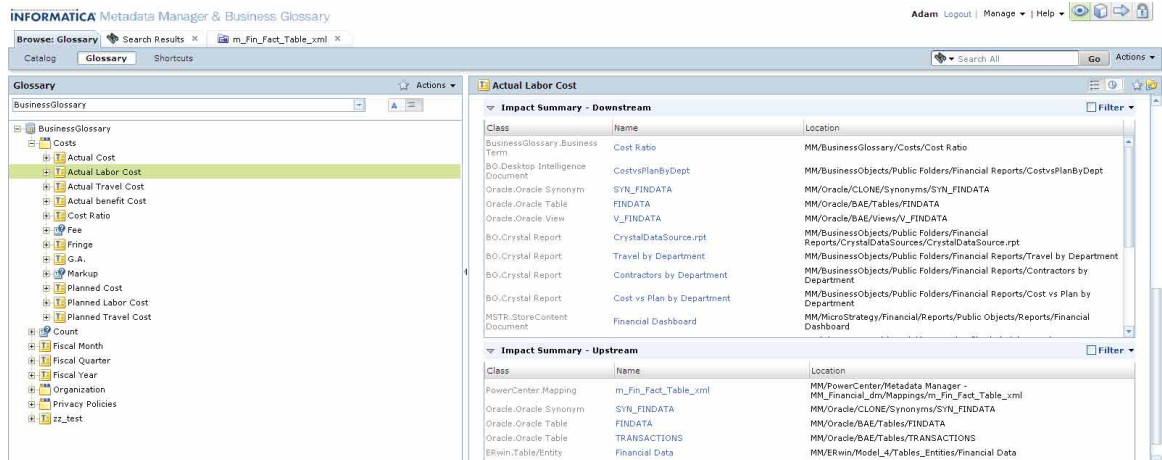


Figure 3: Metadata Manager automatically calculates an impact analysis for every data object selected in the data integration environment.

For people working to ensure the delivery of trustworthy data, this is an essential capability. For example: If a business user asks IT for a change in the way a key performance indicator (KPI) is calculated, what would be the impact? Changing the calculation is relatively easy. The hard part of the analysis is to understand which other data objects in a potentially huge data integration environment depend on that KPI being calculated exactly the way it is calculated today. Finding and dealing with these cross dependencies can take a very long time.

Of course, these capabilities deliver productivity savings. One financial services customer estimated impact analysis saves it more than 95 percent of the effort that goes into the analysis phase of implementing a change. Similarly, a healthcare company shared that impact analysis has saved it an average two weeks of work in planning changes.

The larger benefits are harder to quantify. What would be the cost of a bad business decision made because a key manager had bad data in a report or analysis? Customers using Metadata Manager & Business Glossary in support of a financial data warehouse have shared that a bad business decision could have had a cost impact of millions of dollars.

The Business Glossary Interface

For business stakeholders participating in a data governance initiative, Business Glossary provides a business-friendly user interface to manage business metadata (see Figure 4).

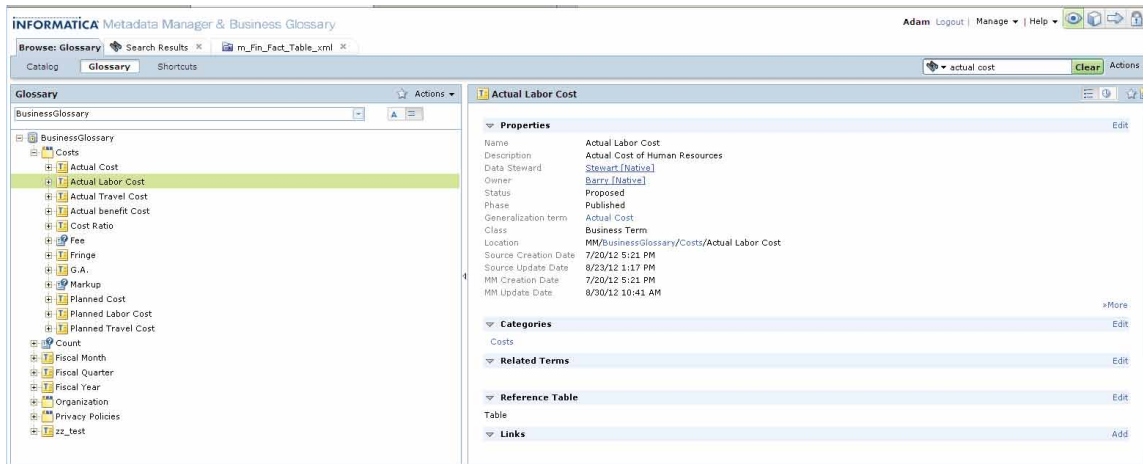


Figure 4: Business Glossary gives business staff and data stewards an interface to create, annotate, and manage business term definitions.

The following capabilities of the Business Glossary interface are key to data governance:

Understanding Business Terminology

Business Glossary provides a central place to define and manage the lifecycle of business terms, definitions, reference data, related terms, links, ad hoc documentation, and notes. A common business vocabulary is essential for clear communications between teams and across business units. It also provides an essential business context for communicating and collaborating with IT on projects.

As an example, one customer shared that its organization had five different currency conversion calculators. Because all of them had legitimate uses, it became critical to understand when to use which one when performing currency conversions. Using the wrong calculator to produce results for an important business decision could lead to incorrect decisions that would have a significant negative impact on the corporation. To prevent this, the organization clearly documented the usage scenarios relevant to each calculator through the Business Glossary interface with a name, definition, and owner of each calculator, and a description of when to use it. Business Glossary in this case became an important tool for defining the customer's conversion calculators.

Assigning Accountability for Business Critical Data

A central goal of data governance is to recognize and assign accountability for data quality and security. Your data governance policy must identify the person or people responsible for the definition, management, usage, and documentation of a business term. The Business Glossary interface offers fields for defining the data owner and data steward. Additional documentation can be included ad hoc.

Ensuring Auditability and Transparency for Compliance

Another key deliverable of a data governance initiative is to provide visibility into who made which changes and when. This visibility supplies an essential audit trail and accountability for these changes and their impact (see Figure 5). In the case of regulatory compliance, this audit trail is often mandatory.

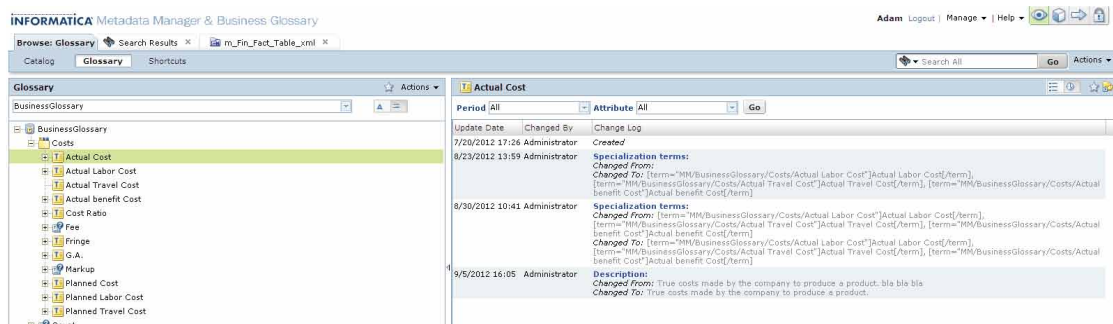


Figure 5: The Business Glossary interface provides an audit trail for changes made to each business term.

The big business benefit from a business glossary is that it improves the quality and speed of communication between business and IT and across business units. The result is greater agility and productivity, fewer errors due to misunderstandings, and less iteration between business and IT on requirements. Ultimately, this means trusted data is the basis of business decisions and processes.

One question we often hear from customers is: "Why can't we just use a wiki to manage our business vocabulary?" The answer is that while wikis have the advantage of being flexible and easy to use, they have several key drawbacks. First, wikis are usually free-form, but in a data governance scenario, you would most likely want to make mandatory the completion of certain fields documenting a business term. For example, a data owner would need to be specified for accountability. Second, wikis have varying levels of capability in terms of tracking who changed what and when. This tracking must be formalized in a data governance or regulatory compliance environment. Finally, the real power of metadata comes when you link the business terms (business metadata) to the underlying technical metadata. That link provides the built-in capability for clear and unambiguous communication between business users and IT staff.

This is a good example of the need to have clearly defined strategies, policies, goals, and requirements as you select tools for data governance.

Practical Considerations for Metadata Management

Best Practices

Many of our clients request a list of best practices about where to begin and how to succeed with metadata management and data governance. Following are some points of advice we've gleaned from conversations with customers who have successfully implemented these initiatives.

1. **Show up, start small, and execute.** Showing up is half the battle. Choose a small project in which you know you can succeed and show tangible results quickly. Then grow from there.
2. **Quantify everything.** Be ready to quantify your results at any time. You never know when you will be asked to justify the existence of your metadata project. Because the ultimate benefit statement—good business decisions, more efficient business processes, and optimized stakeholder interactions based on good data—is so far removed from IT, it is often hard to quantify. Limit the scope of your effort by identifying the critical few business processes, decisions, and interactions that will benefit from these efforts and work on quantifying only those, while building a qualitative business case for all others. As an example, what would be the cost of a bad investment decision because of bad data used in an investment model delivered through your data warehouse?
3. **Set a focused and reasonable scope.** Many metadata and data governance efforts fail because they attempted to accomplish too much at once. Scope-creep and analysis paralysis quickly overwhelmed the projects, which ultimately were cancelled when they failed to produce any meaningful results. No one has the time or budget for endless meetings that fail to produce tangible business results.

“Don't win the Nobel Prize. Keep metadata management simple and use it.”

– Ralph Kimball

4. **Get executive sponsorship.** Your project won't succeed without the right executive sponsorship. Don't shortchange the importance of getting the right sponsorship from both the business and IT sides of the house. Several customers have told us that this is what they did right after their first failure. This support is also critical for getting other groups to contribute to the overall cause.
5. **Establish a data governance initiative.** With a data governance initiative, a company moves beyond metadata as an IT productivity tool and into use cases that have much broader business benefit—and significantly more value—across the organization. An early objective should be the formation of an executive data governance council or steering committee to help set the overall vision and priorities for data-driven opportunities. An overall framework for business-to-business and business-to-IT collaboration is also required to ensure success.

6. Pick a high-value, low-complexity target as your pilot. Choose a specific problem that can return some tangible value, but that won't require boiling the ocean, per se, from the standpoint of data-volume, organizational, or architectural complexity. Then aim to solve the problem before expanding into other areas. For example, provide complete data lineage and full business terms and definitions around a new data warehouse or master data management implementation. The important thing is to pick a project that has high value and high strategic importance to the overall business, but won't get crushed by the weight of its complexity while you're still maturing your data governance processes. That first win is critical. (See also point #2.)

7. Assign owners and get business users involved. Data without business context has no value.

A table with a column name of NW_Net_Revenue has no particular meaning unless it is attached to objects such as:

- A business term
- A term definition
- A term owner
- Reference data (Does "NW" refer to a geographic region? What is in that region?)
- Other documentation and comments

Once you have this business context, you can link the business terms and definitions to the underlying technical metadata, creating a common language between business and IT that will improve communication and collaboration. One key component of this is to be sure to assign owners to the data and metadata. Accountability is critical.

8. Use both the carrot and the stick. Think about the "carrot"—creating incentives, such as:

- Better and easier access to business term definitions, encouraging business users to participate in content creation because they benefit from the results
- Contests for contribution, yielding rewards for contributing new content
- Contents for definitions (using gaming technology), encouraging friendly competition across stewards and motivate them to create good content
- Visibility, perhaps in the form of a simple ranking or recognition of the highest contributors of new (and good) content

It is also good to think about the "stick." Having the backing of executive management and a strong data governance council can go a long way toward ensuring support and buy-in as well. But it's important to use both approaches. All "stick" can be heavy-handed and result in only grudging support, while all "carrot" might feel optional and with limited repercussions for not prioritizing participation.

The challenge is to get the business community to *want* to participate. In very few cases is it the full-time job of people on the business side to provide this context; they have other responsibilities and are simply doing this "on the side." The question is: What's in it for them? You will have to find incentives. The best practice overall is to ensure that participation is formally documented in employee performance management objectives. This will both incentivize them to participate and preclude them from distraction when the inevitable fire drill arises.

9. Tie metadata management to a business initiative. Two customer companies shared that they each tried and failed at data governance initiatives—twice. The new approach they both took was not a top-down initiative, but instead attached metadata management and data stewardship best practices to important new projects as they came through IT. This approach is much more pragmatic, scoped for success, and likely to succeed. It is also more likely to show quick and measurable results.

10. Look for a potential data crisis and be ready for when it inevitably happens. What if your management simply won't fund a metadata management or data stewardship initiative? As a senior manager at a financial services company advised, "Just watch for a data crisis," and be ready with your proposal in-hand. A crisis will happen. What matters is that you are prepared with a proposal that shows how to prevent this type of crisis in the future.

Although these 10 points of advice won't guarantee the success of your metadata or data governance initiative, they will go a long way toward improving your odds of succeeding.

Sizing Your Metadata Effort for Data Governance

One question we frequently hear is, "How many resources should I dedicate to the metadata management part of my data governance initiative?" There is no rule of thumb here. The size of the metadata management staff will vary in relation to such variables as:

- The number of business terms being managed
- The number of topic areas you have for business terms. (These could be by business units, product lines, functional groups, etc. A topic area will require a data steward and some data owners/subject matter experts.)
- The extensiveness of your data integration environment. (How many data sources, applications, data warehouses/data marts, reports, etc. are within the scope?)

Here are three of examples of sizing and growth that we have seen at customer sites:

1. A large bank began implementing Metadata Manager & Business Glossary for Basel II compliance. It began with a single person. Four years and a major acquisition later, it has a staff of 13 people on the project. The cost of the initiative hasn't been questioned because its value has been so clear to the company.
2. A large investment company began with technical metadata management. Soon thereafter, it built one common business glossary to consolidate seven different business glossaries. The business benefit from that project was so clear that other business units within the company also requested business glossaries for their areas. The company has had to dedicate additional data stewards and data owners for each of the business units, but the overall growth of the IT staff has been very small because Informatica Metadata Manager & Business Glossary has provided the automation to scale up their support with minimal incremental head count.
3. A high technology company *attempted three times* to complete a data governance initiative. Because it failed to show value soon after launch during the first two attempts, people drifted away from the project. On the third try, it decided to implement data governance and metadata management as best practices for any new business initiative. It hired a VP of data governance and assigned business staff to fill the role of data owners but has not significantly increased the size of the IT staff to date.

Other Important Tools for Data Governance

This paper is focused on metadata management, but it is important to consider other tools that contribute to a holistic data governance initiative. Some examples of these tools include:

- **Data Quality**, with capabilities in data discovery, data profiling, the creation and management of rules and workflows that remediate data quality issues, and measurement and monitoring of data quality on an ongoing basis.
- **Information Lifecycle Management**, enabling the management of data through its lifecycle, with capabilities for managing:
 - Data archiving and retention
 - Data subsets
 - Data masking for privacy
- **Master Data Management**, for the management and delivery of authoritative data relating to key business processes, such as a customer, vendor, or product. These views are created from data collected across multiple systems into a single “golden record.”

Each of these tools provides key capabilities that contribute to data governance.

Conclusion

Data governance initiatives can become very large and complex. Because of this tendency, many of them fail to deliver quantifiable business results in a timely manner. A successful approach therefore requires taking important preliminary steps to prepare an initiative. First, define the strategy, goals, and scope carefully. Then, define the policies, processes, and people skills that will drive your initiative. And, above all, focus on governing areas that have the greatest business and strategic value to the company. Only then will you be in a position to define the specific requirements for metadata management tools to support the data governance initiative.

Once the groundwork is laid, look for tools that provide the degree of control you need but that also have the flexibility to deal with new and changing requirements. It is important that these tools recognize the different usage requirements of business and technical users and offer specific user interfaces for each. At the same time, they must ensure the ability to foster communication and productive collaboration between these different users in the management and delivery of information to support your business objectives.

Informatica provides a versatile and robust product with precisely these capabilities to deliver trustworthy, secure data and launch successful metadata management initiatives. Its Metadata Manager & Business Glossary offers unique benefits that enable IT managers to minimize the business risk of damaging critical business data while implementing changes. This integrated metadata management product can help your IT organization to:

- Increase efficiency by delivering trustworthy data in support of business decisions and processes
- Boost IT productivity by reducing the time required to analyze planned changes
- Decrease project delivery time by accelerating and enhancing the collaboration between business and IT

With its ability to streamline the handling and delivery of information to support business objectives, Informatica Metadata Manager & Business Glossary can be an essential component in achieving sound data governance.

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.



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

© 2013 Informatica Corporation. All rights reserved. Informatica® and Put potential to work™ 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.