TDWI CHECKLIST REPORT

Cost Justification for Metadata Management

By Philip Russom

2 FOREWORD

2 NUMBER ONE
Recognize that metadata glues enterprises together.

3 NUMBER TWO
Justify metadata management by its benefits.

4 NUMBER THREE
Reduce development time with reusable metadata.

5 NUMBER FOUR
Depend on a centralized metadata repository.

6 NUMBER FIVE
Support many applications with metadata management.

6 NUMBER SIX
Improve reporting via deeper metadata management.

7 NUMBER SEVEN
Extend metadata usage to increase its value.

8 NUMBER EIGHT
Know and evangelize the many justifications for metadata management.

9 ABOUT OUR SPONSOR

9 ABOUT THE AUTHOR

9 ABOUT TDWI RESEARCH
Many people have trouble understanding why metadata management should be executed to exacting standards—or even executed at all. This is natural, because metadata—as managed by most organizations today—has been kept invisible to all but data specialists and a few business people. Up to this point, it has not been the actual deliverable the way that solutions for data integration, data quality, master data management, and operational applications are. These applications and others wouldn’t be possible, much less successful, without metadata. But the assistance provided by a quality metadata management solution is masked because it is embedded within a larger solution, or it has been a degree or two of separation from the business processes it enables.

Pinning a dollar value or other metric to metadata’s contribution is problematic—but it can be done. Even when the metric is fuzzy, metadata’s influence on IT and business success is evident. And you need metrics to provide a credible business case that justifies further investment in metadata management.

This TDWI Checklist Report explains how to provide justification for time and money invested in metadata management solutions. The report points out a few ways to quantify metadata’s contribution to IT projects and business initiatives. It also discusses numerous benefits and how they further justify the development of a modern metadata management solution.

A first step toward justifying the effort and cost of metadata management is to recognize that it is the glue that binds enterprises together, especially in the four business areas in which metadata is commonly deployed:

1. **Design.** Metadata is a component of most well-designed business processes, data models, and applications. Reusable, shared metadata boosts productivity in design and enables accurate project scoping and impact analysis.

2. **Operation.** Digital support for day-to-day production and commercial activities isn’t possible without metadata’s definitions of information, especially when exchanging data across businesses—whether units of the same company or diverse enterprises.

3. **Management.** Managing operations, tracking corporate performance, making strategic decisions based on facts, and other business intelligence functions gain accuracy, credibility, and auditability from careful metadata management.

4. **Governance.** Metadata helps catalog and identify enterprise data assets that require or would benefit from governance for the sake of compliance, security, privacy, system performance, and partner activities.

From these examples, you can see that metadata and its management are both omnipresent and binding, which makes metadata an inevitable best practice for successful data-driven business processes and technical implementations.
A recent survey by TDWI Research asked respondents to identify the benefits of improving metadata management. There are many bottom-line and top-line benefits, as seen in the responses to the survey (see Figure 1).

**What would be the benefits of improving metadata management?**

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better data sharing across an enterprise</td>
<td>79%</td>
</tr>
<tr>
<td>More compliant use of data</td>
<td>59%</td>
</tr>
<tr>
<td>Faster response to change</td>
<td>52%</td>
</tr>
<tr>
<td>Reduced data-related risk</td>
<td>49%</td>
</tr>
<tr>
<td>More effective data management projects</td>
<td>45%</td>
</tr>
<tr>
<td>Reduced costs for the business</td>
<td>39%</td>
</tr>
<tr>
<td>Smoother operations for the business</td>
<td>32%</td>
</tr>
<tr>
<td>Increased revenue for the business</td>
<td>16%</td>
</tr>
<tr>
<td>Other</td>
<td>5%</td>
</tr>
</tbody>
</table>

Figure 1. Based on 458 responses from 122 respondents. Source: TDWI Technology Survey, May 2010.

**Broad data sharing is the top benefit of good metadata management (79%).** Conversely, the same survey revealed that limited data sharing is the leading problem of poor metadata. Sharing data across diverse business units and their information systems is key to business integration, 360-degree views of business entities (e.g., customers, products), customer service, syncing data across multiple ERP or CRM applications, and so on. Hence, metadata greatly facilitates broad data sharing, which in turn contributes to the success of important business initiatives and processes.

**Metadata management yields more compliance and less risk.** These benefits are apparent in the survey results, where users selected more compliant use of data (59%) and reduced data-related risk (49%). For example, when business rules are embedded in metadata, they help control data access and use, which yields better compliance. When metadata describes data structures clearly, the documentation reduces the risk of users selecting the wrong data for a report, portal, data set, or other application.

**Metadata enables faster response to change (52%).** For example, metadata helps data warehouse professionals assemble analytic data sets quickly to study a recent emergency, such as customer churn, grid outages, runs on products, or missed sales quotas. Likewise, quality metadata gives speed and accuracy to data migrations, application upgrades, and corporate reorganizations.

**Metadata reduces bottom-line costs and complexity.** The next section of this report drills into metadata’s influence on IT productivity and other cost issues, proving that solid metadata management truly leads to more effective data management projects (45%). Similar benefits include reduced costs for the business (39%) and smoother operations for the business (32%).

**Metadata supports top-line revenue growth (16%).** TDWI Research has interviewed users who talked about metadata’s assistance in speeding up changes to applications and data that are required for new product development and rollout, careful targeting for sales and marketing campaigns, and on-boarding new partner companies. These business processes lift revenue, and good metadata management practices give the lift greater speed and accuracy.
NUMBER THREE
REDUCE DEVELOPMENT TIME WITH REUSABLE METADATA.

It’s easy to forget that in most industries, payroll is one of the greatest—if not the greatest—expenses a corporation must bear. TDWI’s position is that metadata management can help reduce payroll costs or (more likely) boost productivity to get more projects out of the current staffing without raising payroll costs. To support this assertion, we can construct a formula for estimating the amount of payroll dollars saved due to productivity gains driven by modern practices in metadata management.¹

\[
\text{Money saved (in payroll dollars) via metadata management} = (h_1 - h_2)dps
\]

Where:

- \( h_1 \) = Hours spent by technical staff collecting and documenting source metadata
- \( h_2 \) = Hours spent extracting the same information from a metadata management solution
  (An assumption is that \( h_2 \) will be a small fraction of \( h_1 \))
- \( d \) = Average number of developers on a project
- \( p \) = Average hourly pay of developers, in dollars or other currency
- \( s \) = Number of applications or data management solutions that draw from the metadata repository

As an example of the formula in action, consider that an experienced technical worker in a two-week period (80 hours) can discover and thoroughly document several useful data sources and targets. Once this foundational work has yielded reusable metadata, the same worker or others can get what they need for a new project from the repository in a day or less (eight hours). TDWI surveys have shown that three full-time workers per data management project is fairly common. Assuming that each worker is paid $100,000 per year and works 40 hours a week for 52 weeks a year, the average hourly pay per worker is approximately $48. Finally, let’s say that the repository is used and reused in 10 projects per year. All totaled, the savings in developer time (expressed in payroll dollars) amounts to $103,680 per year.

Saving the equivalent of $100,000 in payroll is great, but what does it mean to an organization? As one of the users interviewed for this report said, “In most cases, metadata management is about efficiency, not cost savings. Time is saved, but the time doesn’t necessarily translate directly into dollars. A more likely translation is into additional work that can be handled by current staffing, which in turn saves money that would’ve been spent on new hires.”

The formula presented here focuses on time spent by various types of developers. It could also apply to other workers who depend on metadata, namely business analysts, data stewards, data governors, report designers, database administrators, and some business people. Alter the formula to include the kinds of people in your organization and plug in numbers that represent their work metrics. Your alteration of the formula should produce a useful indication of potential savings from reusable metadata.

USER STORY
HOW METADATA MANAGEMENT REDUCED PAYROLL COSTS.

“We’re working on an enterprise logical data model (ELDM),” said Ngan MacDonald, who works in information governance and stewardship at Health Care Service Corporation (HCSC). “So far, the ELDM has 2,000 terms that are officially sanctioned as enterprise data references. Each term describes a business entity or attribute, and each is defined via metadata. We make the model and metadata accessible and shareable by managing them with our managed metadata environment.

“When deploying a new billing system recently,” MacDonald continued, “we found that we could reuse 1,200 terms from the ELDM—with minimum tweaking—in a mere eight hours of work. Creating definitions for the other 1,200 terms that the application required took a whopping 292 person hours. This really brought home the real-world reuse and ensuing productivity gains that result from an enterprise data model that’s thoroughly documented through metadata management.

“Projecting from this experience, we can now credibly estimate an average time savings of 90% when applying the ELDM definitions to new applications versus creating new definitions from scratch.”

¹This formula is based on work by Stephen Putnam, senior consultant at Baseline Consulting and a recognized expert in metadata management. Read Steve’s blog posts at www.baselineconsulting.com.
The productivity gains described in the prior Checklist item are impressive. But note that they depend on specific best practices in metadata management. In particular, the reuse of collected and documented metadata depends very much on the use of a metadata repository. Furthermore, the repository must be centralized, in the sense that there’s one repository or a small collection of integrated or federated repositories. The central metadata repository should be governed by a combination of technical and business staff to control access to the repository and create and enforce standards for metadata and other content within the repository.

For some organizations, centralizing metadata management is a big change, because for years they’ve managed metadata in silos, typically with an isolated metadata solution per application, database, or data management tool. For some organizations, successfully centralizing metadata may require the acquisition of an independent metadata repository that’s vendor-agnostic and suited to broad enterprise use, managing metadata for many IT systems from a shared, central source. However, centralizing metadata management is well worth the effort because it leads to a number of benefits relative to cost and best practices.

**Cost benefits.** As seen in the previous Checklist item, a central repository enables the reuse of metadata objects across multiple data management projects, and the increased developer productivity leads to payroll savings or the reallocation of human resources. Other cost benefits may result from having fewer metadata solutions and less complexity to manage.

**Best practices benefits.** A central metadata repository enables many best practices. These include controlled access to metadata (for compliance and security), standards (for consistency in data, metadata, and master data), glossaries (for cataloging data, business definitions, and implementation relationships), and ontologies (for describing relationships across data objects). Furthermore, central repositories tend to support modern technologies that extend the reach and automation of metadata management, such as Web services, service-oriented architecture, matching functions, automated data mapping, and real-time interfaces.

Put those categories of benefits together, and you see that a modern metadata management solution, based on a central repository, has the potential to both reduce costs and enable best practices.

**USER STORY**

**EVERY DOLLAR SPENT ON METADATA SAVES SIX.**

In a recent *ComputerWorld* article, a representative of Intel Corp. explained a best practice the company has developed in metadata management. Instead of relying on tools that extract metadata from multiple systems long after metadata’s origins, their solution includes a process that captures metadata as it’s created or changed. Intel found that this keeps its central, shared repository of metadata more up to date and accurate than other methods. Once all development teams felt that the repository managed “trusted metadata,” they began using it regularly. Based on the productivity gains and greater data consistency that resulted from earnest use of the central metadata repository and reuse of its contents, the Intel representative quoted in the article reported saving $53 million in development costs in one year. He also estimated that every dollar spent on metadata saves $6 in development and IT costs.
When TDWI Research asked users which applications they manage metadata for, a whopping 81% of survey respondents pointed to business intelligence (BI) and data warehousing (DW)—far more than for any other application type (see Figure 2). Next in the pecking order come the data management disciplines data integration (44%), data quality (32%), CDI (26%), and MDM (22%). Numerous operational applications registered in the survey, including those for financials (28%), CRM (27%), ERP (16%), and HR (16%). We often think of metadata as solely for structured data, but a number of survey respondents reminded us that it’s also important in content management and/or Web content management (27%).

**For what kinds of enterprise applications is your organization actively managing metadata?**

<table>
<thead>
<tr>
<th>Application</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business intelligence and/or data warehousing</td>
<td>81%</td>
</tr>
<tr>
<td>Data integration</td>
<td>44%</td>
</tr>
<tr>
<td>Data quality</td>
<td>32%</td>
</tr>
<tr>
<td>Data governance</td>
<td>29%</td>
</tr>
<tr>
<td>Financial applications</td>
<td>28%</td>
</tr>
<tr>
<td>Content management and/or Web content</td>
<td>27%</td>
</tr>
<tr>
<td>Customer relationship management (CRM)</td>
<td>27%</td>
</tr>
<tr>
<td>Customer data integration (CDI)</td>
<td>26%</td>
</tr>
<tr>
<td>Master data management (MDM)</td>
<td>22%</td>
</tr>
<tr>
<td>Data stewardship</td>
<td>19%</td>
</tr>
<tr>
<td>Enterprise resource planning (ERP)</td>
<td>16%</td>
</tr>
<tr>
<td>Human resources applications (HR)</td>
<td>16%</td>
</tr>
<tr>
<td>Other</td>
<td>5%</td>
</tr>
</tbody>
</table>

Figure 2. Based on 462 responses from 124 respondents. Source: TDWI Technology Survey, May 2010.

The survey results illustrate that metadata management serves a broad range of applications and data management practices. In this context, we can consider metadata management to be a supporting practice, in a support role similar to that of data modeling, data profiling, and data glossaries. Supporting practices aren’t as high profile or “sexy” as primary, core practices, such as data integration, data quality, MDM, and popular operational applications. But secondary support practices still provide essential functionality, without which most applications and data management wouldn’t be possible. Yet supporting solutions are so embedded within primary solutions that it’s hard for business people—and some technical personnel, even—to see them and value their contributions appropriately. We all need to be diligent in remembering supporting practices like metadata management when we celebrate the success of primary solutions.

Reporting is a great example of a data management practice that benefits directly from quality metadata management, then passes the benefits directly and visibly to business end users. Furthermore, reporting and its issues are already understood by executives, so reporting may be a good place to start when soliciting their support of metadata management solutions.

Reporting has risen in prominence recently for a variety of reasons, and its success, as always, continues to depend on how metadata driven it is. For example, we now live in the age of accountability. Your organization can be audited at any point by regulatory agencies, business partners, employees, investors, and so on. You typically answer their audit questions with data in reports. They will inevitably demand an audit trail confirming the sources of report data, which well-managed metadata can provide. Hence, to survive an audit, you need to build a credible audit trail for all information—and metadata is a key component of that trail.

Audits aside, BI and DW become higher priorities all the time, and reports are their main product. Metadata-driven reporting makes it easy to answer common questions from report consumers, such as “Where did the data in this report come from?” “How was the data altered to fit the report?” and “How current is the data in this report?” After all, users need answers before making critical decisions, and credible answers based on quality metadata build user confidence in BI and DW as the source of trusted data.

**USER STORY**

**A METADATA-DRIVEN AUDIT TRAIL IS ESSENTIAL TO SURVIVING AN AUDIT.**

A few years ago, this author was working for a small software vendor that was audited by the U.S. Federal Trade Commission (FTC). Initially, the FTC asked for financial reports and operational information, which we quickly delivered. They were happy with the information, but also asked us to show an audit trail for it. We had no audit trail because of the sorry state of metadata management in our operational systems and reporting platform. It took us many months to construct an audit trail. The delay made the FTC suspicious, so they deepened and lengthened the audit. When the audit ended three years later, it had cost us several million dollars in expenses. We estimated that it would have cost about one-third of that if we had had proper metadata management and other facilities for a credible audit trail.
Traditionally, metadata has literally been “data about data,” describing technical attributes such as data type, data structure, and data sources or dependencies. However, metadata isn’t just about the data’s technical attributes anymore. More and more, metadata practices have expanded to describe other attributes, some of them business oriented, such as business definitions and data usage rules. This is driven by the shift toward a knowledge economy that the industry is experiencing globally. Business users of many types need more of a hands-on approach to information, which means that metadata tools need to support business users better by presenting data in business terms.

According to a TDWI survey (see Figure 3), roughly half of users surveyed (49%) follow the traditional approach of describing technical attributes, whereas the other half (51%) also includes business-related information in metadata.

Which of the following best describes the content of metadata in your organization?

- Only describes technical information (e.g., data structures, sources) - 49%
- Also describes business information (e.g., business definitions, usage rules) - 51%

Figure 3. Based on 121 respondents. Source: TDWI Technology Survey, May 2010.

TDWI feels confident that the trend toward extending metadata’s uses will continue as organizations go deeper into governance (which defines rules for data access and usage), compliance (which demands a digital audit trail), and master data management (which defines business entities represented in data). These practices and others benefit from a broad range of information embedded in metadata.

In a similar trend, metadata repositories have been extended to manage much more than metadata. Nowadays, they also store and provide broad access to application routines, project management documents, master and reference data, data quality metrics, reports germane to development, and so on. When centralized, the repository is a conduit through which developers, managers, and other people collaborate. To facilitate collaboration, some repositories support browser-based access, security for multiple user roles, annotations for managed objects, and discussion threads.

Consider the power that today’s knowledge workers could derive from contextual and instantaneous access to the metadata underlying any report, spreadsheet, or document in which they are working. No longer would there be a misunderstanding of data, terms, or the process that produced the data. This powerful “pop-up” capability is enabled by an enterprise metadata repository, and it’s available in the market today.

Considering these modern practices, metadata management’s value has been extended, just as metadata’s documentation and the uses of the repository have. This added value further justifies the deployment of a modern metadata management solution.

(Continued, next page)
(Continued)

USER STORY

BUSINESS METADATA IS KEY TO BUSINESS INVOLVEMENT.

“We failed with a couple of attempts at starting a data warehouse, largely because we couldn't get people to agree on exactly what business entities the data represents, plus the context for each entity and relationships among them,” said the metadata architect at a leading U.S. healthcare provider and insurance company. “Consider a simple definition, like admission date. If the patient entered an outpatient clinic, received some treatment, and was then rushed to the emergency room, when did admission begin? Was it in the clinic or in the emergency room? This might be calculated differently for an insurance claim versus clinical outcome reporting. So, getting the business to accurately define and correlate different-but-related definitions is key to consistent reporting and data warehouse development.

“We all knew that these entity definitions and relations could be determined by and documented through metadata management. But taking a traditional approach, which merely captures system metadata, was too techie and arcane for the business people involved. We started getting a lot more traction when we complemented system metadata with business metadata, which describes entities and their relations using terms that business people know and understand.

“You can start by capturing system metadata, then add business metadata atop it,” continued the metadata architect. “But we got business people really excited and involved when we turned the process around so they can lead it. Today, our business people use a straightforward Web-based tool to specify entities, contexts, and relationships that they need reports about, and then technical personnel find the right data, document metadata properly, and build reports. The application is based on the metaGlossary component of ASG’s Rochade. The approach enables business folks to lead and take control as they define business requirements up front with a hands-on application. Everyone knows the application wouldn’t be possible without business metadata, which has raised the visibility and importance of metadata management in everyone’s minds.”

As we’ve seen in this TDWI CheckList Report, there are numerous credible justifications for metadata management and its costs. These are summarized in the following list:

Justify metadata management by its benefits. The top benefit is broad data sharing, which reminds us that metadata glues together multiple business units and processes. Other benefits include better compliance, less risk, faster response to change, and top-line revenue growth.

Reduce development time with reusable metadata. This is exactly how users quoted in this report saved millions of dollars in IT payroll costs.

Depend on a centralized metadata repository. It’s the secret sauce in most of the benefits and cost justifications discussed here. Sharing and reusing metadata across multiple teams and projects isn’t likely without a vendor-agnostic central repository.

Support many applications with metadata management. Metadata’s supporting role is critical. Always remember this when celebrating the success of the numerous applications it supports.

Improve reporting via deeper metadata management. All reports need an audit trail, and metadata enables the trail. This builds confidence in data warehouse data and speeds up all types of audits.

Extend metadata usage to increase its value. In turn, the higher perceived value further justifies metadata management. Continue to document technical metadata, but extend that into business metadata. Use metadata to support new types of applications, such as master data management and data governance. Foster collaboration via a repository and shared metadata. Embrace advanced practices, such as metadata pop-ups.

Review the above list occasionally and share it with your colleagues and management. The list succinctly reminds everyone that metadata management reduces cost and risk and contributes to the success of many business initiatives and technology implementations.
ABOUT OUR SPONSOR

ASG
tSoftware Solutions

ASG is the industry-leading provider of enterprise metadata management solutions. ASG-Rochade® is a powerful metadata repository that manages information about data and systems across the enterprise, uncovering complex relationships between information technology assets, facilitating what-if analysis, and increasing the productivity of business and IT. ASG-metaGlossary™ is a business glossary solution that provides a governed and contextually rationalized vocabulary between business and IT to help proactively address data governance and compliance mandates. Founded in 1986, ASG provides world-class solutions to more than 85 percent of the world’s largest companies, enabling them to reduce costs, improve business-service delivery, and reduce risks.

www.asg.com

ABOUT THE AUTHOR

Philip Russom is the senior manager of TDWI Research at The Data Warehousing Institute, where he oversees many of TDWI’s research-oriented publications, services, and events. He’s been an industry analyst at Forrester Research, Giga Information Group, and Hurwitz Group, where he researched, wrote, spoke, and consulted about BI issues. Before that, Russom worked in technical and marketing positions for various database vendors. You can reach him at prussom@tdwi.org.

ABOUT TDWI RESEARCH

TDWI Research provides research and advice for BI professionals worldwide. TDWI Research focuses exclusively on BI/DW issues and teams up with industry practitioners to deliver both broad and deep understanding of the business and technical issues surrounding the deployment of business intelligence and data warehousing solutions. TDWI Research offers reports, commentary, and inquiry services via a worldwide Membership program and provides custom research, benchmarking, and strategic planning services to user and vendor organizations.

ABOUT THE TDWI CHECKLIST REPORT SERIES

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.