

Metadata Management:

The Hero in Unleashing Enterprise Data's Value

Connect all the pieces of your data management and data governance lifecycles to reduce risks and realize results

► Introduction

Most organizations don't use all the data at their disposal to reach deeper conclusions about how to drive revenue, achieve regulatory compliance or accomplish other strategic objectives.

They don't know what they don't know, and this problem is only getting worse. The size of the global datasphere is projected to reach 163 ZB by 2025, according to IDC, with all of the disparate data sources in legacy systems, new system deployments, and the creation of data lakes and data warehouses.

Finding, ingesting, integrating, linking, sharing and analyzing data historically has relied on manual processes and therefore takes a lot of time, money and specialized technical resources. So how can enterprises become more adept at wringing all the value they can from their data assets?

The hero in unleashing your data's real potential is metadata management because it answers a lot of important questions, such as:

- What data do we have?
- Where did it come from?
- Where is it now?
- How has it changed since it was originally created or captured?
- Who is authorized to use it and how?
- Is it sensitive or are there any risks associated with it?

Creating and sustaining an enterprise-wide view of and easy access to underlying metadata is a tall order. The numerous data types and data sources that exist today weren't designed to work together, and data infrastructures have been cobbled together over time with disparate technologies, poor documentation and little thought for downstream integration. Thus the applications and initiatives that depend on a solid data infrastructure may be compromised, rendering faulty analyses and insights.

Pinpointing what data exists and where goes hand in hand with being able to understand what it all means within the context of the business and in common terms that are adopted across the enterprise. Having that consistency is the only way to ensure insights generated by analyses are useful and actionable, regardless of business department or user exploring a question. In addition, policies, processes and tools that define and control access to data by roles and across workflows are critical for security purposes.



► Introduction

continued

When approached manually, it is expensive, time-consuming, error-prone and can't keep pace with a dynamic enterprise data management infrastructure. The reality is there's not enough time, people and money for true data management using manual processes. Your organization won't be able to take complete advantage of analytics tools to become data-driven unless you establish a foundation for agile and complete data management.

METADATA MANAGEMENT HELPS YOU



Discover data. Identify and interrogate metadata from various data management silos.



Harvest data. Automate the collection of metadata from various data management silos and consolidate it into a single source.



Structure and deploy data sources. Connect physical metadata to specific data models, business terms, definitions and reusable design standards.



Analyze metadata. Understand how data relates to the business and what attributes it has.



Map data flows. Identify where to integrate data and track how it moves and transforms.



Govern data. Develop a governance model to manage standards, policies and best practices and associate them with physical assets.



Socialize data. Empower stakeholders to see data in one place and in the context of their roles.

METADATA MANAGEMENT DEFINED

Quite simply, **metadata** is data about data. It's generated every time data is captured at a source, accessed by users, moved through an organization, integrated or augmented with other data from other sources, profiled, cleansed and analyzed. Metadata is valuable because it provides information about the attributes of data elements that can be used to guide strategic and operational decision-making.

Metadata management is the administration of data that describes other data, with an emphasis on associations and lineage. It involves establishing policies and processes to ensure information can be integrated, accessed, shared, linked, analyzed and maintained across an organization.

► The Role of Metadata Management

A strong data management strategy and supporting technology enables the data quality the business requires, including data cataloging (integration of data sets from various sources), mapping, versioning, business rules and glossaries maintenance and metadata management (associations and lineage).

A comprehensive data governance strategy plus technology will enable an organization to determine master data sets, discover the impact of potential glossary changes across the enterprise, audit and score adherence to rules, discover risks, and appropriately and cost-effectively apply security to data flows, as well as publish data to people/roles in ways that are meaningful to them.

An accurate perspective of the corporate metadata landscape reduces friction in data accessibility and utility, improves overall quality, and accelerates digital transformation as more individuals from across the organization become adept at reporting and data analysis.

This real-time, high-quality picture is only possible with metadata management tools that can identify any data from anywhere, including “data at rest” in databases, data warehouses and data lakes and “data in motion” as it is integrated with and used by key applications, and control the metadata landscape to facilitate collaboration and limit risk.

When evaluating metadata management tools, the following capabilities are essential in building an automated, real-time, high-quality data pipeline so all stakeholders – data scientists, data stewards, ETL developers, enterprise architects, business analysts, compliance officers, CDOs and CEOs – can access the relevant data they’re authorized to use to guide their actions in producing the desired results.

- Reference data management for capturing and harmonizing shared reference data domains
- Data profiling for data assessment, metadata discovery and data validation
- Data quality management for data validation and assurance
- Data mapping management to capture the data flows, reconstruct data pipelines, and visualize data lineage
- Data lineage to support impact analysis
- Data pipeline automation to help develop and implement new data pipelines
- Data cataloging to capture object metadata for identified data assets
- Data discovery facilitated via a shared environment allowing data consumers to understand the use of data from a wide array of sources

Integrating and automating data management and data governance is still a new concept for many organizations, but the advantages are clear.

Better data quality. With automation, data quality is systemically assured with the data pipeline seamlessly governed and operationalized to the benefit of all stakeholders. Data issues and inconsistencies within integrated data sources or targets are identified in real time to improve overall data quality by increasing time to insights and/or repair. It’s easier to map, move and test data for regular maintenance of existing structures, movement from legacy systems to new systems during a merger or acquisition or a modernization effort.

► The Role of Metadata Management

continued

Quicker project delivery. Automated enterprise metadata management provides greater accuracy and up to 70 percent acceleration in project delivery for data movement and/or deployment projects. Harvest metadata from various data sources and map any data element from source to target and harmonize data integration across platforms. With this accurate picture of your metadata landscape, you can accelerate Big Data deployments, Data Vaults, data warehouse modernization, cloud migration, etc.

Faster speed to insights. High-paid knowledge workers like data scientists spend up to 80 percent of their time finding and understanding source data and resolving errors or inconsistencies, rather than analyzing it for real value. That equation can be reversed with stronger data operations and analytics leading to insights more quickly, with access/connectivity to underlying metadata and its lineage. Technical resources are free to concentrate on the highest-value projects, while business analysts, data architects, ETL developers, testers and project managers can collaborate more easily for faster decision-making.

Greater productivity & reduced costs. Being able to rely on automated and repeatable metadata management processes results in greater productivity. For example, one erwin DI customer has experienced a steep improvement in productivity – more than 85 percent – because manually intensive and complex coding efforts have been automated and 70+ percent because of seamless access to and visibility of all metadata, including end-to-end lineage. Significant data design and conversion savings, up to 50 percent and 70 percent respectively, also are possible with data mapping costs going down as much as 80 percent.

Regulatory compliance. Regulations such as the General Data Protection Regulation (GDPR), Health Insurance and Portability Accountability Act (HIPAA), Basel Committee on Banking Supervision (BCBS) and The California Consumer Privacy Act (CCPA) particularly affect sectors such as finance, retail, healthcare and pharmaceutical/life sciences. When key data isn't discovered, harvested, cataloged, defined and standardized as part of integration processes, audits may be flawed. Sensitive data is automatically tagged, its lineage automatically documented, and its flows depicted so that it is easily found and its use across workflows easily traced.

Digital transformation. Knowing what data exists and its value potential promotes digital transformation by 1) improving digital experiences because you understand how the organization interacts with and supports customers, 2) enhancing digital operations because data preparation and analysis projects happen faster, 3) driving digital innovation because data can be used to deliver new products and services, and 4) building digital ecosystems because organizations need to establish platforms and partnerships to scale and grow.

An enterprise data governance experience. Stakeholders include both IT and business users in collaborative relationships, so that makes data governance everyone's business. Modern, strategic data governance must be an ongoing initiative, and it requires everyone from executives on down to rethink their data duties and assume new levels of cooperation and accountability. With business data stakeholders driving alignment between data governance and strategic enterprise goals and IT handling the technical mechanics of data management, the door opens to finding, trusting and using data to effectively meet any organizational objective.

► How to Do Metadata Management

erwin's metadata management software, the erwin Data Intelligence Suite (erwin DI), harmonizes data management and data governance processes in an automated flow.

Organizations then have visibility and control over their disparate data streams – from harvesting to aggregation and integration, including transformation with complete upstream and downstream lineage and all the associated documentation.

erwin DI is unique in its ability to automatically harvest, transform and feed metadata from a wide array of data sources, operational processes, business applications and data models into a central data catalog and then make it accessible and understandable within the context of role-based views.

Our ability to integrate and continuously refresh metadata from an organization's entire data ecosystem, including business processes, enterprise architecture and data architecture, forms the foundation for enterprise-wide data discovery, literacy, governance and socialization.

With an automated, real-time, high-quality data pipeline, enterprise stakeholders can make strategic decisions on a full inventory of reliable information. Applications and initiatives that depend on a solid data infrastructure are based on sound analysis and insights.

erwin DI builds on our diverse industry expertise and proven data modeling software, which also generates large volumes of technical metadata. It also integrates with third-party data access rights management and security tools. We've worked with high-profile brands in financial services, healthcare, technology and critical infrastructure, as well as government agencies, for more than 30 years. Now many of these customers are using erwin DI to help them confront complex issues in highly sophisticated and regulatory environments to support ERP systems, data warehouses, data lakes, big data projects, cloud databases and even graph solutions.

erwin DI sits on a common metamodel that is open, extensible and comes with a full set of APIs. A comprehensive list of erwin-owned standard data connectors is included for automated harvesting, refreshing and version-controlled metadata management. Optional erwin Smart Data Connectors are capable of reverse-engineering ETL code of all types and connecting bi-directionally with reporting and other ecosystem tools. These connectors offer the fastest and most accurate path to data lineage, impact analysis and other detailed graphical relationships.



► How to Do Metadata Management

continued

Additionally, erwin DI is part of the larger erwin EDGE platform that integrates data modeling, enterprise architecture, business process modeling, data cataloging and data literacy for data intelligence. We know our customers need an active metadata approach to:



Understand their business, technology and data architectures and the relationships between them



Create an automate a curated enterprise data catalog, complete with physical assets, data models, data movement, data quality and on-demand lineage



Active their metadata with integrated business glossaries and data dictionaries that provide business context for stakeholder data literacy

And we work with our customers to co-develop the solutions in the erwin EDGE, ensuring we deliver products they can rely on and adjusting our roadmap based on their input.

erwin has been recognized as a Leader in the Gartner “2019 Magic Quadrant for Metadata Management Solutions,” which provides a detailed overview of the market, including evaluations of 17 vendors based on completeness of vision and ability to execute.



**THE REPORT AND THE QUADRANT GRAPHIC CAN BE
DOWNLOADED HERE: www.erwin.com/GartnerMMMQLoader.**



► erwin DI's Key Metadata Management Features

The erwin DI Suite combines erwin's data catalog and data literacy capabilities so enterprise stakeholders have greater awareness of and access to available data assets, guidance on their use, and guardrails to ensure data policies and best practices are followed.

erwin Data Catalog (erwin DC) automates enterprise metadata management, data mapping, code generation and data lineage for data movement and/or integration and modernization of the data architecture. The solution harvests metadata from a broad variety of data sources and maps data elements from source to target, including “data in motion,” while harmonizing data integration across platforms. erwin DC includes:

erwin Metadata Manager. Creation and maintenance of a sustainable metadata foundation for data preparation, management, governance and consumption, automating manual tasks to increase efficiencies, quality and time to value for data development and deployment

erwin Mapping Manager. An integrated development environment for creating and maintaining source-to-target mapping and transformation specifications to centrally control data movement, integration and transformation

erwin Reference Data Manager. Management of reference data of all types to provide visibility and control, reduce costs and ensure accuracy and consistency of use across enterprise systems

erwin Lifecycle Manager. Control for migration, modernization and integration projects across the system development lifecycle to reduce development time and costs, while improving visibility and accuracy

erwin Data Quality. Integrated data profiling along with data quality assessment for registered data assets and their associated metadata to reduce the data discovery and assessment cycle and ensure high-quality results, with ease of use for any role

erwin Data Literacy (erwin DL) enables data stewards to curate and govern data assets so data consumers can discover data relevant to their roles and understand it within a business context. By creating data communities and promoting data fluency, stakeholders have an integrated and contextual view of the semantic, business and technical aspects of the entire data landscape. erwin DL includes:

erwin Business Glossary Manager. Gives data assets business context (terminology, policies, rules) to ensure business visibility, understanding and alignment, thus lowering governance costs, reducing data discovery and analysis cycles, while dramatically accelerating the time to meaningful insights

Business User Portal. Web-based, self-service portal geared primarily toward business users to facilitate easy access and collaboration around data management, consumption and governance

erwin DI customers may opt to add **erwin Smart Data Connectors**, automation and developer productivity tools that harvest data-in-motion directly from a broad variety of code types and industry standard tools and languages (BI, ELT, ETL). These tools provide visibility, improve processes and integrate data consumption platforms with reusable code, saving considerable time and money.

► erwin DI in Action

Our customers are hard at work to address enterprise metadata management challenges, and we'd like to highlight one particular case study demonstrating the power of metadata management and data intelligence in a highly regulated industry.

E.ON is a privately-owned energy supplier based in Essen, Germany, with thousands of employees throughout the United States, Europe and Asia focused on energy networks, customer solutions and renewables.

The company has become a modern pioneer within the energy sector and invests in new ideas and products to generate, store and optimize energy efficiency. For example, it has installed more than one million smart energy meters that give customers more control of their energy usage and costs with convenient, automatic payment options.

Part of E.ON's evolution entails digitizing the vast amounts of data generated by such large customer and employee bases. The challenge is managing that data and turning it into meaningful information to help both the business and customers.

To accomplish such digital transformation mission, the company has implemented a data governance/intelligence program. E.ON is treating data as a strategic asset in shaping products and services in addition to using it to better understand customers and ensure compliance with health, safety, environmental and data security and privacy standards.

The company selected erwin as its data governance partner because of its comprehensive metadata management, data harvesting and automated data lineage capabilities. As a result, E.ON will be able to serve the right data to the right people according to their roles and responsibilities, with confidence in its quality. Then they can use it to make better, faster decisions that ultimately improve the energy company's performance.

"When we started digitalizing the company, we wanted to put a system of data governance in place to identify data assets, simplify documentation, and improve the quality of information throughout our organization. Our goal is to make the most of our data assets to generate value for the company."

Juan Bernabé-Moreno
Chief Data Officer
E.ON



► erwin DI in Action

continued

erwin DI has been used to deliver “data governance as a service” for E.ON’s one-stop data access and lineage within the organization’s data lake that includes numerous sources such as ERP, CRM, BI, ticketing and customer service systems. The goals of this initiative are to standardize the company’s vocabulary, meet regulatory compliance requirements, and understand and then optimize how data is used across global operations.

Since the go-live, the erwin software has provided E.ON with more transparency and enabled a paradigm shift from application-centric to data-centric processes to fuel the use of artificial intelligence (AI). Various E.ON subsidiaries now can see and understand their data assets, because of properly documented data ownership, a commonly agreed-upon business language to describe data, and separation of data domains.

In addition, erwin DI shows the flow and transformation of data assets, while enabling data quality issue diagnosis and backtracking, in migration and ongoing steady-state data integration scenarios. These capabilities prove especially helpful in establishing semantic mappings between data systems.

The central data governance team has begun developing a framework to quantify the value delivered to the organization, considering both efficiency increases and incremental AI-delivered value. Data scientists and AI professionals in particular benefit from the transparency and the shorter “time to data,” which, at the scale of E.ON, results in several million euros of additional revenue.

Be sure to review the entire spectrum of a vendor’s offerings. Some may boast of a strong business glossary, for example, but lack functionality for maintaining business rules and processes. Others may not provide data quality features that specifically target compliance requirements. Some may boast of automation but actually require outside professional services to generate lineage. But you can have it all by using a single solution that checks every box.

It’s time to get smarter about metadata management. erwin DI will help you connect your data management and data governance efforts, automating and accelerating your speed to meaningful insights and better decision-making.



REQUEST A DEMO OF erwin DI.



About erwin, Inc.

As the data governance company, erwin provides enterprise modeling and data intelligence software to help customers discover, understand, govern and socialize their data to mitigate risks and realize results. The erwin EDGE platform facilitates IT and business collaboration in driving actionable insights, agile innovation, risk management and business transformation. We help government agencies, financial institutions, healthcare companies and other enterprises around the world unlock their potential by maximizing the security, quality and value of their data assets.

Connect with
erwin at erwin.com

