



Sponsored by: **erwin**

Authors:
Stewart Bond
Matthew Marden

September 2020

The Value of Robust Data Intelligence to Enable Data Governance with erwin

IDC OPINION

For an organization to be successful in this new digital economy, it must be data-driven and have a good data culture, meaning it requires technology and people that can manage and understand data from capture to consumption. According to the online Merriam-Webster dictionary, data is defined as “factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation.” But without intelligence, data is just data. The future of intelligence is informed by data intelligence, where data-native workers and machines can synthesize information, improve their capacity to learn, and automate insights at scale.

IDC defines data intelligence as business, technical, relational, and operational metadata that provides transparency of data profiles, classification, quality, location, context, and lineage, providing people, processes, and technology with trustworthy, reliable data. Data intelligence enables the enterprise disciplines of data governance and data quality management and improves data literacy and integrity within an organization so that data self-service is achievable. This data intelligence is at the heart of the erwin Inc.’s Data Governance solutions, helping organizations achieve a future of intelligence in the digital economy.

In this research, IDC spoke with organizations that have deployed the erwin Data Intelligence Suite (erwin and erwin DI) for data governance to support their growing data environments and open possibilities for harnessing the power of data intelligence to fuel their businesses. Like many organizations, the erwin customers we interviewed are balancing needing to ensure the security, integrity, and quality of their data with the business imperative to generate data-driven insights and value. They uniformly described erwin as a fundamental component of their efforts to generate more value from data while minimizing data-related risk. While these study participants were at different points in their adoption cycles of the erwin DI Suite at the time of their interviews, they spoke to common benefits and anticipated gains related to use of erwin that included:

- **Higher data quality, integrity, and availability** due to automated data mapping and validation capabilities
- **Improved data governance and stewardship** as the result of cross-organization data visibility, robust metadata capabilities, and the ability to enforce data-related rules and policies
- **Broader employee access** to data and insights through self-service functionality
- **More robust analytical activities** as the result of improved awareness, understanding, and access to data
- **Business gains and employee productivity improvements** by finding new ways to generate value through data

SITUATION OVERVIEW

Data is the lifeblood of the digital economy — it is what is driving new business models, better customer experiences, better decision-making, and artificially intelligent automation. The global pandemic in 2020 has accelerated digital transformation and amplified the value of data in what will become the next normal as the global economy struggles through recovery. In a world where market conditions, supply chains, work locations, and communication methods are constantly changing, data is a constant source that can be used to inform decisions from crisis to recovery. To use data effectively, it needs to be trusted, understood, and used appropriately, and herein lies many problems that organizations face in the digital economy.

The amount of data that is being generated in the digital economy is staggering, with IDC estimating 45ZB of data created in 2019. This figure is expected to grow at a CAGR of 26% over five years to 2024 (see *Worldwide Global DataSphere Forecast, 2020–2024: The COVID-19 Data Bump and the Future of Data Growth*, IDC #US44797920, April 2020). IDC has also predicted a bump in the amount of data that is being generated because of the pandemic. The results of an IDC data integration and integrity (DII) end-user survey completed in 2019 also shows that data is diverse, with 95% of organizations integrating up to six different types of data across 10 different types of data management technologies as they manage operations, seal strategic insights, and make business decisions. The move to the cloud has also been accelerated as secure access to data and computing resources is required from anywhere work is being done. In 2019, 94% of organizations were integrating data across hybrid cloud environments according to the results of the survey referenced previously.



Data intelligence is the key to unlocking the value of data, improving the efficiency and effectiveness of data-native workers, improving the level of trust that organizations have in data, and enabling the organization with data through governance.

The diversity, distribution, and scale of data in the digital economy are challenging data enablement in the organization. Data quality and trust issues surface because the location, source, and provenance of data are unknown; data context and meaning are unknown or misunderstood, impacting data literacy efforts; and accountability for data and metadata is absent. If everybody is responsible for the data, nobody is responsible for the data.

Effects of data environment complexity and the state of intelligence about data are being seen in the efficiency and effectiveness of data-native workers. The 80/20 rule stating the percentage of time spent in data discovery and preparation compared with the percentage of time spent in analytics is getting worse, now approaching 85/15 as per the results of IDC's 2019 DII survey. This survey also told us that on average, data-native workers are more unsuccessful than successful in their tasks as they search for, prepare, govern, and analyze data.

Data intelligence is the key to unlocking the value of data, improving the efficiency and effectiveness of data-native workers, improving the level of trust that organizations have in data, and enabling the organization with data through governance. Knowing where data is; what data means; who is accountable and responsible for data; who uses data, when, how, and why; and how dirty or clean data is and discovering inherent relationships within the data itself bring clarity of understanding of complex data environments. The scale, distribution, and diversity of modern data environments also demand automation, including harvesting metadata and analysis of metadata using assisted and unassisted machine learning for autoclassification, semantic mapping, relationship and anomaly detection, and policy creation and application.

Each of the organizations interviewed in this research exhibited many of the symptoms of data environment complexities, and each was able to articulate qualifiable and quantifiable benefits and expected benefits of data intelligence delivered through the erwin data governance solutions to improve data enablement, data literacy, and data protection efficiencies and the effectiveness of data-native workers and machines that learn from the data.

ERWIN DATA GOVERNANCE SOLUTIONS

The erwin Enterprise Data Governance Experience (EDGE) platform is enabling data governance, literacy, and engineering solutions for organizations competing in the digital economy. The erwin DI Suite, part of the EDGE platform, is delivering intelligence about data to data-native workers and machines, providing business and technical data context to ensure the right data is being used by the right resources, at the best time, and for the right reasons.

The erwin Data Intelligence (erwin DI) suite is the heart of the erwin EDGE platform, as it combines data catalog and data literacy capabilities with built-in automation for greater awareness of and access to available data assets, guidance for asset use, and guardrails to ensure data policies and best practices are followed.



The erwin DI Suite consists of two modules: erwin Data Catalog (erwin DC) and erwin Data Literacy (erwin DL). The erwin Data Catalog product automates enterprise metadata management, data mapping, reference data management, sensitive data identification, code generation, data lineage, and impact analysis for faster time to insight and value. The erwin DL product includes a business glossary and self-service data discovery tools that enable data understanding, curation, association, and governance in business context to improve collaboration, analysis, and decision-making for data-native workers. erwin DI combines data catalog and data literacy capabilities to deliver technical, business, and semantic intelligence of the organization's data universe.

erwin DI also has a set of optional erwin Smart Data Connectors that can be leveraged for advanced automation and access to metadata stored across many data and analytic assets such as ETL/ELT platforms, business intelligence (BI) reports and dashboards, stored database procedures, automated testing tools, ecosystem utilities, and ERP applications. erwin is the same company known for the erwin Data Modeler that has been in the market for decades, and that data modeler plus capabilities in business process and enterprise architecture modeling can be added to the solution to expand understanding and knowledge of data in full business and technical context.

VALUE OF THE ERWIN DI SUITE

Firmographics

IDC spoke with four organizations about their use of the erwin DI Suite. Interviews were designed to understand its impact on various aspects of their data management and governance efforts. Interviewed organizations were headquartered in the United States and included:

- A financial services organization that provides financing to enterprise customers
- A healthcare provider focused on pharmacy and care delivery services
- A retailer selling pet supplies
- A utilities company serving a major metropolitan city/region

Data-Related Challenges Facing Interviewed Organizations

Interviewed erwin customers described challenges common to organizations that must effectively reconcile explosive growth in the volume, complexity, and criticality of data with the need to ensure data quality, governance, stewardship, and access. They understood the importance of establishing and enforcing common norms and policies for data definitions and data use across their organizations. However, they acknowledged that siloed data ownership and not having a common platform for data governance have hampered their efforts. As a result, they have faced various friction points as they attempt to generate more value from operational data, which have made it more challenging to establish and enforce data-related rules and policies. Study participants provided details about these challenges and how they link to their decision to use the erwin DI Suite:

- **Need for data literacy in terms of common definitions and understanding:** *"The main challenge that erwin helps with is data literacy (i.e., being able to find the data across the organization, understand it, and have common definitions of different metrics). Today, if there are two different data lakes or two different departments, they will probably give two different answers about a data-related question. By using erwin, we hope they will be able to get to the same exact definition of the metric and then give the same answer."*
- **Single platform for collecting and accessing cross-organizational data:** *"A key part of data management is governance. Before erwin, there was no single point where all information was collected — most of our information was in Word documents."*
- **Need for validation of source data:** *"What we were looking for with erwin was a solution that could validate our data across client data coming from multiple data sources, where we could write business and technical rules on it."*
- **Need to minimize complexity:** *"The user-friendliness of erwin was most important to us ... We used [another vendor solution] on another project when we were trying to develop a common data model; it's a good tool, but in retrospect, it would have been too difficult."*

“ The user-friendliness of erwin was most important to us ... We used [another vendor solution] on another project when we were trying to develop a common data model; it's a good tool, but in retrospect, it would have been too difficult. **”**

Choice of the erwin DI Suite

Study participants described various criteria for choosing erwin, but their choices boiled down to its functionality and cost. The limitations of their tools and approaches in terms of generating value through data had become increasingly evident. They struggled to implement robust data governance rules and ensure data quality and integrity without a common platform, and they were unable to convince many line-of-business (LOB) users to take

“One of the things that we liked about erwin was it had good visual representation of how data could be connected.”



With erwin, interviewed organizations not only reduce data-related friction and risk but enable analytics and other employees who rely on data to better apply data and insights to support businesses.

ownership of and become stewards of their data. As a result, data too often fell into silos that obstructed cross-organizational use, and LOB users found it challenging to create value and competitive differentiation through data.

Interviewed erwin customers provided more details about their specific considerations:

- **Ability to provide users visual, graphical representations of data:** “One of the things that we liked about erwin was it had good visual representation of how data could be connected. This helps users understand that, and the primary reason we were looking at erwin was really for data literacy across the organization.”
- **Visual outputs, ability to flag sensitive data:** “The graphical lineage and being able to click on the end product to provide a visual output of the data source are key features of erwin for us. The second important factor is the capability to flag source tables or source applications related to sensitive information. If used properly, this tool can be used to flag critical information.”
- **Cost and functionality:** “erwin proved out certain use cases which I knew were not feasible using [another solution] out of the box. You needed another tool along with it for the lineage ... The other factor why we went with erwin was cost. [Other solutions] are multiples of the cost that erwin was ready to provide.”
- **Mapping and data validation functionality, relationship with erwin:** “What really was important to us when we chose erwin was finding a mapping manager and a data validation tool ... erwin also gave us metadata management and data lineage. Also, the salespeople and everybody we dealt with were very helpful ... We felt really comfortable with erwin. They knew the product, and they were there to help.”

Assessing the Value of the erwin DI Suite

IDC’s research shows the value and potential value for organizations of using the erwin DI Suite platform to improve their data management and governance efforts through robust automation, improved data definitions, and lineage capabilities and having a cross-organizational data platform. As a result, interviewed organizations not only reduce data-related friction and risk but enable analytics and other employees who rely on data to better apply data and insights to support businesses.

“ *The benefit we’re going to have with erwin is updated mapping. Everything we did before was manual. Every time we validated data, we had five to six people validating who would write their own queries. We’re standardizing our validation rules with erwin’s support.* **”**



Several interviewed erwin customers noted that being able to create robust business glossaries and data dictionaries with erwin allows for establishment of common data definitions.

Higher Data Quality and Integrity

Interviewed organizations’ ability to achieve higher value through data depends to a significant extent on the quality and integrity of their data. They explained that erwin has either already helped or they expect it to enable substantial improvements in these areas. Specifically, they stressed the importance of moving from manual data mapping processes to standardized and automated processes. Further, they cited erwin’s strong metadata functionality. One study participant commented on mapping-related benefits: *“The benefit we’re going to have with erwin is updated mapping. Everything we did before was manual. Every time we validated data, we had five to six people validating who would write their own queries. We’re standardizing our validation rules with erwin’s support.”*

Study participants also cited improved capabilities with erwin related to data ingestion and lineage as important differentiators. Without a common platform, they found that lack of transparency or even mistakes or errors impeded these activities. With erwin, they have automated or are automating these efforts, thereby reducing the number of manual touch points, minimizing the likelihood of errors, and maintaining more robust data. One interviewed customer explained: *“With erwin, we have automated scheduled metadata ingestion. All databases are connected, and we ingest data on a weekly or monthly schedule. We are also moving toward automated lineage ... which was completely manual.”*

Study participants spoke to various staff and operational benefits related to automation of data mapping, ingestion, and lineage, including higher data quality and enhanced data governance foundations. Importantly, they noted clear gains in terms of data accessibility. For example, two interviewed organizations reported that users now require 90%+ less time to locate data (92% and 95% faster with erwin). This increased speed to identification of data reflects how erwin has enhanced the flow of data across their organizations and underscores the potential impact in terms of changing how employees and businesses can access needed data to drive business activities.

Improved Data Governance and Stewardship

Study participants also stressed the centrality of erwin in implementing and enforcing data governance and stewardship. They acknowledged challenges in these areas related to fast data growth and entrenched and inconsistent practices in terms of maintaining, organizing, and using data. Such inconsistencies can substantially impede organizations’ efforts to leverage data in support of their businesses and increase operational risk related to data loss or breach.

Interviewed erwin customers described how erwin either has already promoted or will promote efforts to improve data governance and stewardship. Several noted that being able

“erwin’s functionality in multiple facets of data governance was important to us, such as the business glossary and data dictionary/ metadata repository, and data lineage as well. erwin allowed us to set up a data governance policy and link our data assets to that.”



erwin DI automatically harvests, transforms and feeds metadata from a wide array of data sources, operational processes, business applications and data models into a central catalog. This catalog is accessible and understandable via role-based, contextual views so stakeholders can make strategic decisions based on accurate insights.

to create robust business glossaries and data dictionaries with erwin allows for establishment of common data definitions. Another organization described how creating “data crosswalks” with erwin would allow for better understanding of the logic behind data mapping.

Interviewed erwin customers provided specifics:

- **Business glossary and metadata repository:** *“erwin’s functionality in multiple facets of data governance was important to us, such as the business glossary and data dictionary/ metadata repository, and data lineage as well. erwin allowed us to set up a data governance policy and link our data assets to that.”*
- **Ability to create data crosswalks for mapping visibility:** *“In erwin, we’re able to crosswalk to all data mapping, including for specific field lanes. Then employees will be able to look back and see any rule or logic or transformation that was applied in the mapping and what mappings those were, all the way back to the data source.”*

For interviewed erwin customers, these functionalities not only enable more robust data governance but can create a foundation for establishing robust data stewardship at an organization level. One study participant explained: *“We’ve never even been able to have data stewardship because we’ve done everything manually. Now, with erwin, we’re going to have stewards over certain sets of data.”*

Study participants provided several specific examples of efficiencies and improvements related to their use of erwin in terms of data governance and stewardship activities, including:

- 24% more efficient metadata mapping
- 30% more efficient data governance and stewardship team

Broader Access to Data

In terms of data use, study participants attributed broader access to data and insights — especially for LOB users — to use of erwin. In particular, they noted that the automation and visibility achieved with erwin enable self-service access to data. Thus LOB users are less dependent on IT teams to deliver the data and insights upon which they rely. One study participant explained: *“Previously, everything was manual, and there was very little self-service. Our business teams couldn’t do their own analysis and always had to rely on IT teams, but IT only has a limited capacity to scale.”* Another commented on its broader vision for using erwin to empower business users: *“The vision we have with erwin is building a data lake concept where we are trying to bring in all data from different business applications and then provide an environment in which employees can do self-service analytics using that data.”*

“One important thing for us was really being able to integrate erwin with other existing systems that we have on our data analytics platform . . . A lot of the other vendors didn't have the ability to integrate, but erwin built it quickly and showed us the integration capabilities to our BI tool.”

Study participants noted several specific examples of efficiencies and improvements related to their use of erwin in terms of data access, including:

- 100% of users having self-service access to data
- 10% efficiencies for IT teams supporting data management and governance activities due to self-service functionality

More Robust Analytical Activities

Study participants want to empower their businesses with data, which means having analytics capabilities that are robust, relevant, and timely. Interviewed erwin customers described how they are leveraging the erwin DI Suite to improve their analytics capabilities, including through integration with other tools and delivering higher-quality data to be transformed into insights:

- **Ability to integrate with BI tool:** *“One important thing for us was really being able to integrate erwin with other existing systems that we have on our data analytics platform . . . A lot of the other vendors didn't have the ability to integrate, but erwin built it quickly and showed us the integration capabilities to our BI tool.”*
- **Provide better understanding of data to support analytical work:** *“erwin should primarily help with understanding how data is organized in our data warehouse and data lake. Once our people understand it, [they] might be able to use it better for their analysis work.”*

Like most organizations, study participants have teams dedicated to creating value through data, including business analysts, data scientists, and analytics engineers. They need these teams to deliver insightful and differentiated analytical work, which requires timely access to data of the highest quality and usability. When this occurs, these teams can create more value for their organizations. For example, one interviewed organization cited a 30% boost in productivity for its business analyst team through its use of erwin.

Business and Employee Productivity Gains

Study participants spoke to both general ways in which erwin has enabled their business activities and its more specific impact on their development and certain LOB teams.

For example, several interviewed erwin customers noted that their development teams benefit from automated scripts and enhanced data lineage functionality. One study participant commented: *“Before erwin, our ETL developers had to write scripts for us on our tables, if they had time to do them. We were very dependent upon those developers to do things for us. Now, we can do it in the erwin tool ourselves and that's amazing.”* Another explained: *“With erwin's reverse lineage*

“We are going to open a business portal up to everyone with erwin. . . We feel that if they can see the metadata and the mapping we’ve done, look at the client information, and look at the glossary, that’s going to help everybody have a better understanding of our business.”

capabilities, our developers accelerate development activities. They don’t have to rediscover things every time they are going in coding and can slice off about 15% on their development time.”

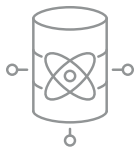
LOB teams can also see significant benefits from enhanced flow of high-quality and robust data with erwin. Increasingly, these teams rely on data to enhance the value they provide to their organizations. For example, one study participant commented on how its finance team has benefited from use of erwin: *“Our accountants and CFO can now see where the data comes from that feeds into their financial reporting, and that has improved their confidence.”*

More generally, study participants linked their use of erwin to broader business enablement. Again, this relates to having the capability with erwin to deliver transparent, high-quality, and robust data from a single platform across their organizations, without running into roadblocks related to data stewardship or governance. Several interviewed erwin customers discussed this in more detail:

- **Creating more value through robust data:** *“As we build out a data glossary with erwin, we can build our terms and our business rules and associate these back to data . . . Now, we spend 80% of our time cleaning up junk. We don’t really get to use the smart people we have to look at this data for what else we could get out of it, because we’re too busy trying to get through an implementation. That’s where I think erwin is going to help us.”*
- **Offer visibility and transparency into data and business activities:** *“We are going to open a business portal up to everyone. We will open to our ETL developers and other staff and our business partners that look at our data. We feel that if they can see the metadata and the mapping we’ve done, look at the client information, and look at the glossary, that’s going to help everybody have a better understanding of our business.”*

CHALLENGES AND OPPORTUNITIES

erwin faces the challenge of discretionary spending choices that are being made during a time of crisis in the economy. While the erwin data governance solutions are key to data enablement, which is also a key to digital transformation that has been accelerated by the global pandemic, the value and return on investment (ROI) will need to be clearly understood for erwin to be part of customers’ digital transformation spending. Data governance solutions by erwin will also need to deliver on the value and meet or exceed ROI to retain customer spend as the economy returns to growth into the next normal. To the extent that customers can realize the benefits and value exhibited in this study, erwin is well positioned to face the economic slowdown.



Interviewed erwin customers described their use of the DI Suite as central to their efforts to derive more value from data. Interviewed organizations were at different stages in their adoption and use of the erwin DI Suite, but they spoke to important gains in enhancing data governance and stewardship, improving understanding of metadata and data lineage, and opening up new opportunities for their businesses.

The pandemic has also created more opportunities for making data intelligence available through software because remote workers no longer have the water cooler or cubicle conversations about where the best data for a specific project or analytic model exists. There is an increased use of external data to help organizations track and predict the economic, societal, and healthcare impacts of the virus. erwin Data governance solutions can help data-native workers better understand external data in context for higher levels of trust, utilization, and new insights analytics of internal and external data blending to manage business continuity and resiliency through the economic crisis.

CONCLUSION

Successful organizations are, by necessity, data-driven with good data cultures. This fundamentally requires them to be able to manage and understand data. To create value through data, organizations must leverage intelligence about data to enable a future of intelligence, including the organization's ability to synthesize information, improve its capacity to learn, and deliver insights at scale. The erwin DI Suite is designed to enable organizations to establish this type of data intelligence and position them to reap the benefits of intelligence in the digital economy.

Interviewed erwin customers described their use of the DI Suite as central to their efforts to derive more value from data. They understand that creating value through data is a multifaceted effort that requires ensuring data quality, integrity, and availability, as well as improving analytics and LOB teams' ability to access and leverage data. While interviewed organizations were at different stages in their adoption and use of the erwin DI Suite, they spoke to important gains that are being made, including enhanced data governance and stewardship, improved understanding of metadata and data lineage, and opening up new opportunities for their businesses to more fully exploit operational data that can be used to enhance decision-making and improve productivity.

APPENDIX

Methodology

IDC's standard Business Value methodology was utilized for this project. This methodology is based on gathering data from organizations using the erwin DI Suite as the foundation for the model. IDC collected quantitative benefit information during the interviews using a before-and-after assessment of the impact of using the erwin DI Suite.

Note: All quantitative results discussed in this study are individual organization-level results and do not reflect an aggregate model of the impact of use of the erwin DI Suite.

IDC Research, Inc.

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2020 IDC.

Reproduction without written permission is completely forbidden.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.