

erwin® Data Catalog (DC)

AT A GLANCE

erwin DC automates enterprise metadata management, data mapping, code generation and data lineage for faster time to value and greater accuracy for data movement and/or deployment projects. Harvest metadata from various data sources and map data elements from source to target, including “data in motion,” and harmonize data integration across platforms. When combined with erwin Data Literacy (DL), you can see an accurate picture of your business data assets for the data intelligence to govern and accelerate digital transformation projects, including Big Data deployments, Data Vaults, data warehouse modernization, cloud migration, etc., without heavy reliance on technical resources.

KEY BENEFITS/RESULTS

- Demonstrates faster time to value and greater accuracy for data movement and/or deployment projects through automated metadata enrichment, data quality metrics, reuse and guided workflows
- Reduces the need for costly manual intervention by automating everything from the profiling of data, the generation of Big Data, ETL s and database procedural code to detailed data lineage and impact analysis views
- Provides a central governance framework for data quality deployment and integration independent of siloed data management technologies
- Supports IT audits and regulatory compliance by governing enterprise data assets as well as data management processes and infrastructure for both IT and business stakeholders

KEY FEATURES

- Accelerated pre-ETL mapping with the drag-and-drop approach, creating mapping specifications and eliminating human errors
- Automated metadata harvesting and cataloging using our wide range of data source and data movement connectors
- Version mapping specifications for tracking changes instantly, as well as integrated requirements, testing and deployment management for the development lifecycle
- Easy import of legacy mappings, plus share and reuse mappings and transformations
- Built-in reference data management including code sets and “cross walks”
- Plug-in code generation for Data Vault and leading ETL/ELT tools, Big Data systems and scripts
- Built-in data profiling and integrated data quality metrics
- Integration with erwin DL for business glossary management and data discovery self-service

SOLUTION OVERVIEW

erwin DC provides an automated metadata-driven approach to enterprise data management and data preparation through these features:

- **Easy data mapping** — Use drag-and- drop and auto-mapping features to accelerate the creation of metadata-driven mapping documents, eliminating manual work and costly errors seen with the traditional Excel approach.
- **Impact analysis** — Instantly identify the impact of change to an attribute or table across the warehouse in a matter of seconds, saving valuable time and resources.
- **Automated metadata harvesting and cataloging** — Configure and schedule regular scans to harvest detailed technical metadata — from the mainframe to Big Data to business intelligence solutions — to establish a strong metadata backbone on which to build data management and data governance practices.
 - Data stores: relational, NoSQL, Big Data
 - File connectors/cloud environments/packaged applications
 - Data modeling tools

#	Column Name	DQ Score	Column Datatype	Length	Locked?	Job State	Total Rows	Distinct Values	% Distinct Values	Repeated Values	Nulls	% Nulls	Min Value	Max Value
1	First_Name	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
2	Last_Name	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
3	Gender	—	varchar	50	🔒	COMPLETED	50008	36	0%	15	0	0%		
4	Street_Address	—	varchar	50	🔒	COMPLETED	50008	49018	98%	1140	0	0%		West
5	City	—	varchar	50	🔒	COMPLETED	50008	3033	6%	1503	0	0%		
6	State	—	varchar	50	🔒	COMPLETED	50008	174	0%	54	0	0%		
7	Zip	—	varchar	50	🔒	COMPLETED	50008	2887	6%	2864	0	0%		
8	County	—	varchar	50	🔒	COMPLETED	50008	1127	2%	558	0	0%	0	Y
9	Country	—	varchar	50	🔒	COMPLETED	50008	90	0%	45	0	0%		
10	Phone_Number	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
11	Social_Security_N	—	varchar	50	🔒	COMPLETED	50008	49978	100%	15	0	0%		
12	Email	—	varchar	50	🔒	COMPLETED	50008	48966	98%	1245	0	0%	CAMWME_ABEE tybol	
13	DOB	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
14	Education	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
15	Company	—	varchar	50	🔒		0	0	0%	0	0	0%	0	
16	Occupation	—	varchar	50	🔒		0	0	0%	0	0	0%	0	

A
DQ Score

16
Total Columns

9
Profiling Columns

50008
Total Rows

49978
Unique Values

0
Nulls

2864
Repeated Values

erwin DC provides built-in data profiling capabilities and quality metrics to augment the technical and business metadata auto-harvested and cataloged from your data management infrastructure.

- **Upstream and downstream lineage** — Generate end-to-end lineage between repositories and view data flows from source systems to the reporting layers, including all intermediate transformation and business logic so stakeholders have information to make swift decisions.
- **Reference data management** — Centrally map and manage reference data code sets to ensure accuracy and quality across enterprise datastores and applications. Versioning and change management capabilities ensure that code sets are applied consistently and their content and usage is easily auditable and transparent to stakeholders.
- **Requirements, testing and deployment management and integration** — Take control of the software delivery lifecycle, while providing visibility to stakeholders with built-in lifecycle management capabilities and integrations.
- **Version management and change control** — With built-in versioning, baseline and archive all mapping documentation and view change comparison reports over time to enable efficient IT and regulatory audits for movement and change-control procedures.
- **Easy adoption for data migration and data movement** — Easily scan metadata from various operational systems for a seamless fit wherever a case for data migration arises. Integration with an in-house ETL tool generates jobs to facilitate the movement of data to the target systems.
- **SAML 2.0 integration** — Accelerate the deployment and adoption of erwin DC by aligning it with your organization’s security and access-control infrastructure.
- **Robust data profiling** — Profile data sources and integrate data quality details and metrics along side metadata details for easy visibility and understanding.
- **Integration with erwin DL** — Benefit from a data intelligence software suite for the management and integration of a business glossary (terminology/policy) with detailed physical metadata through a role-based portal that enables stakeholders to discover and understand enterprise data within context.

CRITICAL DIFFERENTIATORS

- **Centralized metadata-driven mapping with support for most ETL and data asset technologies** — Efficiently transform and move data in tight alignment with business-driven requirements from a single, unified platform.
- **Enterprise metadata catalog** — Automate the harvesting, cataloging and visualization of any metadata from anywhere, including reference data, data movement processes, data lineage and impact analysis.
- **Lifecycle management** — Manage and socialize the gathering of functional requirements, mapping design, testing and release management of enterprise data movement and transformation projects.
- **Built-in reference data management, data profiling and data quality** — Ensure the context of your data is understood, under control and visible to all stakeholders.

For more information, please visit www.erwin.com.

As the data governance company, erwin provides enterprise modeling, data cataloging and data literacy 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.