

EDM Webinar

Accelerate Cloud Data Platform Adoption with Automated Data Intelligence

A conversation with



Danny Sandwell
Director of Product Marketing
erwin



Moderated by **Mike Meriton**

Co-Founder & COO, EDM Council

- Joined EDM Council full-time 2015 to lead Industry Engagement
- EDM Council Co-Founder & First Chairman (2005-2007)
- EDM Council Finance Board Chair (2007-2015)
- Former CEO GoldenSource (2002-2015)
- Former Executive for D&B Software and Oracle
- FinTech Innovation Lab – Executive Mentor (2011 – Present)



Danny Sandwell

Director of Product Marketing, erwin

- An IT industry veteran focused on delivering value from data for more than 30 years.
- Responsible for communicating the technical capabilities and business value of the company's data intelligence solutions.
- During Danny's 20+ years with the erwin brand, he also has worked in pre-sales consulting, product management, business development and business strategy roles
- His goal is to help enterprises unlock the value of their data assets while mitigating data-related risks.



The Business Drivers for Modern Data Architecture



Perpetual
Transformation



Data-Driven
Innovation



Data
Agility



Risk
Mitigation



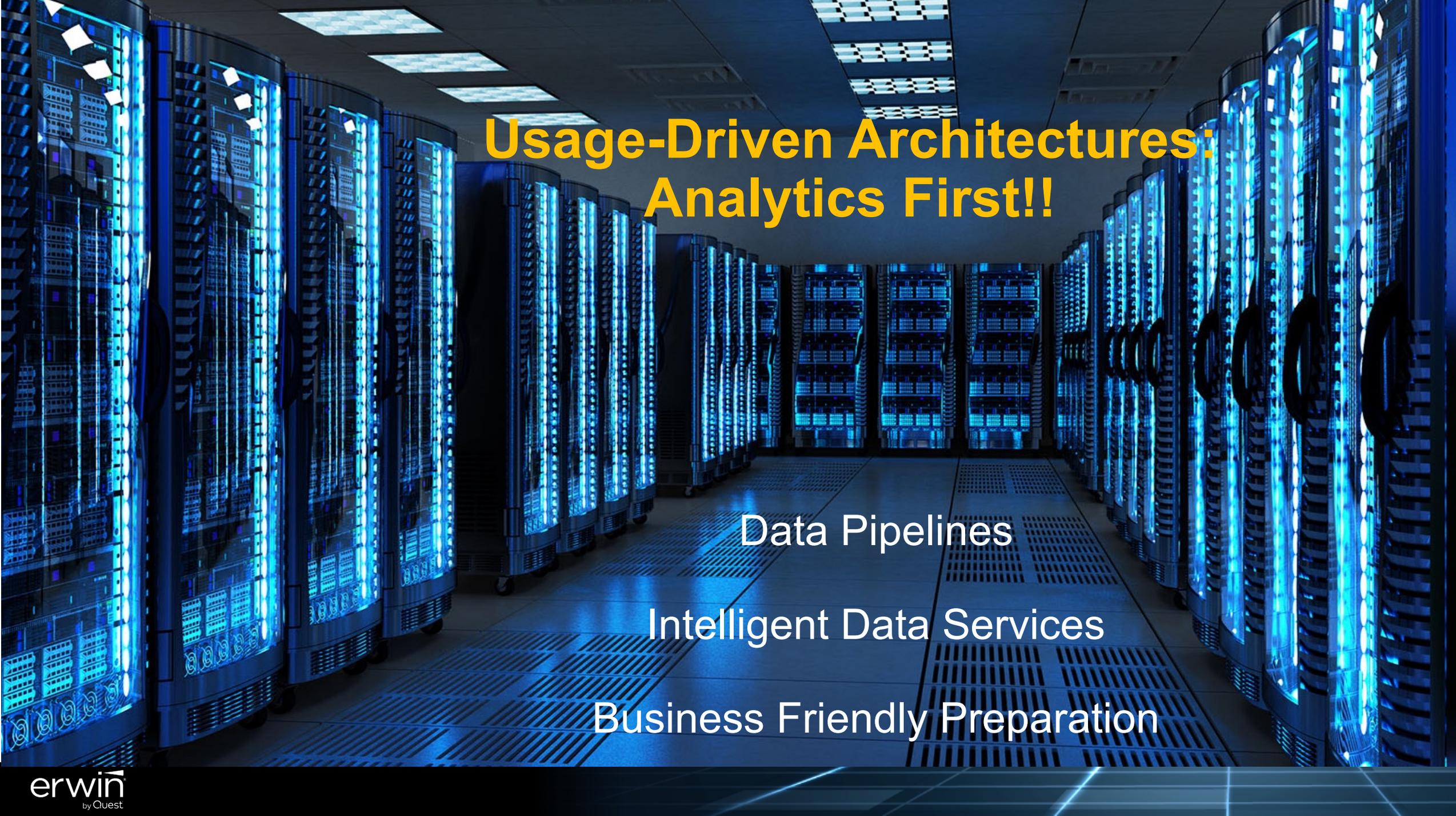
Why Modernize? The Enterprise Data Dilemma

➤ 95% of organizations integrating at least six different types of data across 10 different data management technologies.

➤ 94% of organizations were integrating data across hybrid cloud environments.

➤ 85% of time spent on data prep and only 15% in analysis hampers data-driven digital transformation.

➤ 50% of organizations lack sufficient AI and data literacy skills to achieve maximum business value



Usage-Driven Architectures: Analytics First!!

Data Pipelines

Intelligent Data Services

Business Friendly Preparation

Smart Data Architectures: Intelligent and Learning Everyday...

Active Metadata

Dev/Ops Automation

Predictive Data Voyages

“From The Cloud Out” Data Architectures Agility, Scale-Ability and Resilience

Strategic Applications

Strategic Workloads

Strategic Platforms



Data Architectures Governed For The Good of All: Mitigating Risk and Assuring Opportunities

Intelligent Governance Frameworks

Data Literacy Facilitation

Built-in Stakeholder Participation



Cloud Data Platform Benefits and Capabilities

Benefits

- Performance and Scalability
- Elasticity and Agility
- Lower TCO and Future Proof
- More Value From Data

Capabilities

- High Performance Data Store
- Hybrid DBMS Modalities
- Agile Data Integration
- Integrated BI & Analytics



Challenges To Realizing Modernization Benefits



Migrating Legacy Deployments

Time To Value

Conversion
Accuracy

Cost
Containment



Data Governance and Intelligence

Migration
Transparency

Documenting
cutting edge
technologies

Data
democratization
enablers

Utilize Your Enterprise Architecture to Plan The Move

1. Business Case

- Drivers
- Risk
- Cost
- Scope
- Benefit
- Plan

2. Architecture Assessment

- Current State
- Priorities
- Complexity
- Requirements
- Impacts
- Security

3. Migration Strategy

- Stakeholders
- Roadmap
- Approach
- Capabilities
- Priorities
- Resources

6. Testing and Operationalization

- Testing
- Support
- Data Quality
- Security
- Monitoring
- Governance

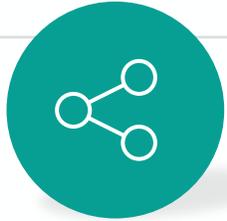
5. Migration



4. Technology Selection

- Future State
- Architecture
- Technologies
- Vendors
- Tools
- Integrators

Data Intelligence Enables Automation and Governance



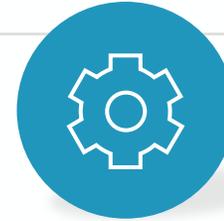
Harvest



Curate



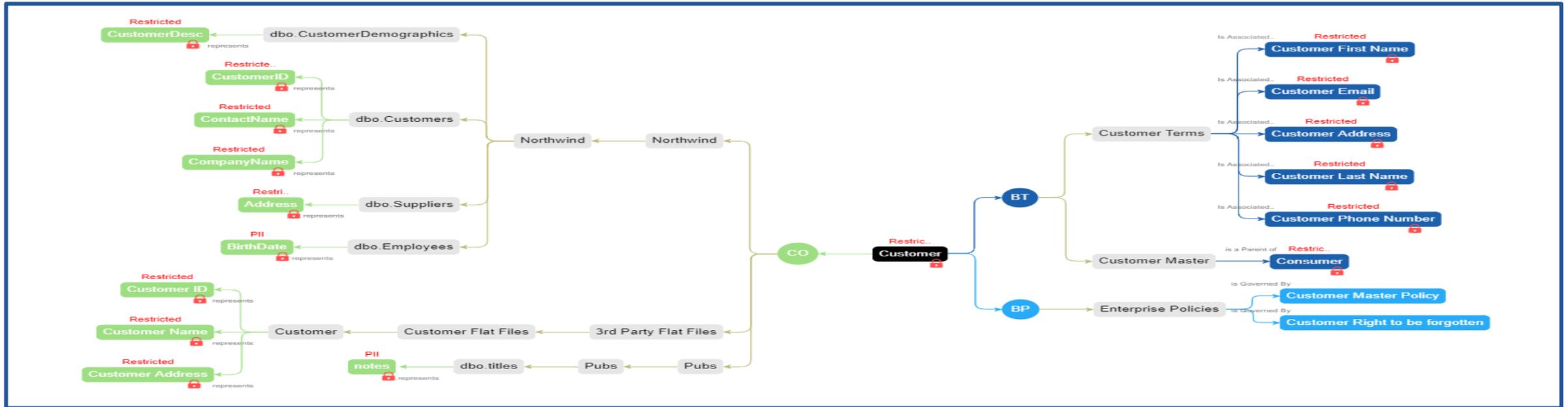
Govern



Activate



Socialize



Data Intelligence: Get More Utility From Your Metadata

Auto Document

- ✓ Data Sources
- ✓ Data Models
- ✓ Data Movement Processes
- ✓ Data Consumption

Auto Curate

- ✓ Technical Asset Associations
- ✓ Business Asset Associations
- ✓ Sensitive Data Classification

Auto Render

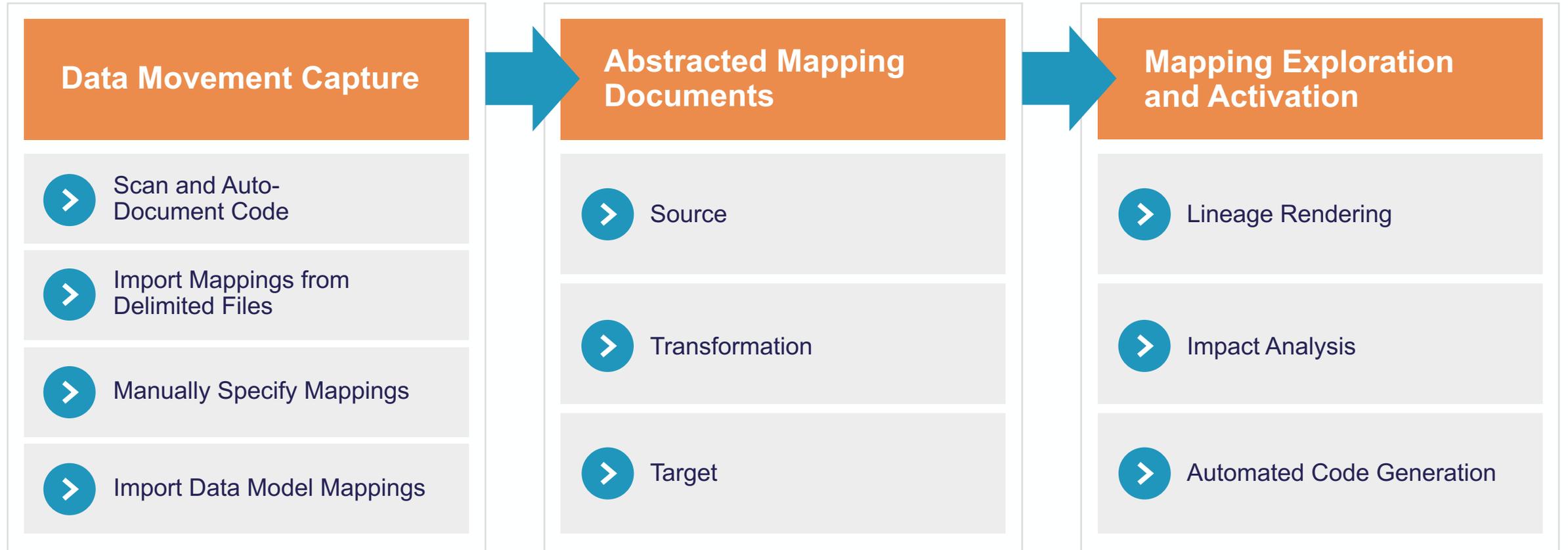
- ✓ End-to-End Lineage
- ✓ Impact Analysis
- ✓ Mind Map Graph Views
- ✓ Focused Dashboards

Auto Generate

- ✓ Data Pipelines
- ✓ Data Workloads
- ✓ Data Movement Code
- ✓ Platform Orchestration

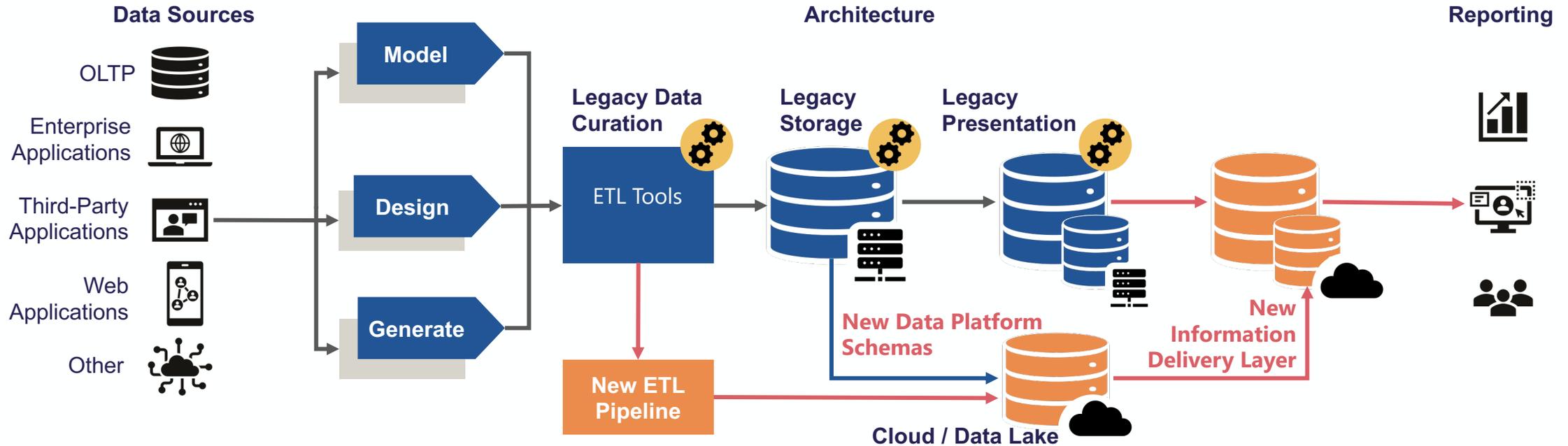


Data Mapping Documents: Activating Metadata For Maximum Utility



Accelerating and Assuring Architectural Modernization

Automated documentation, transformation, code generation and governance



Convert Data Structures

Accelerate Data Migration

Transform Data Movement

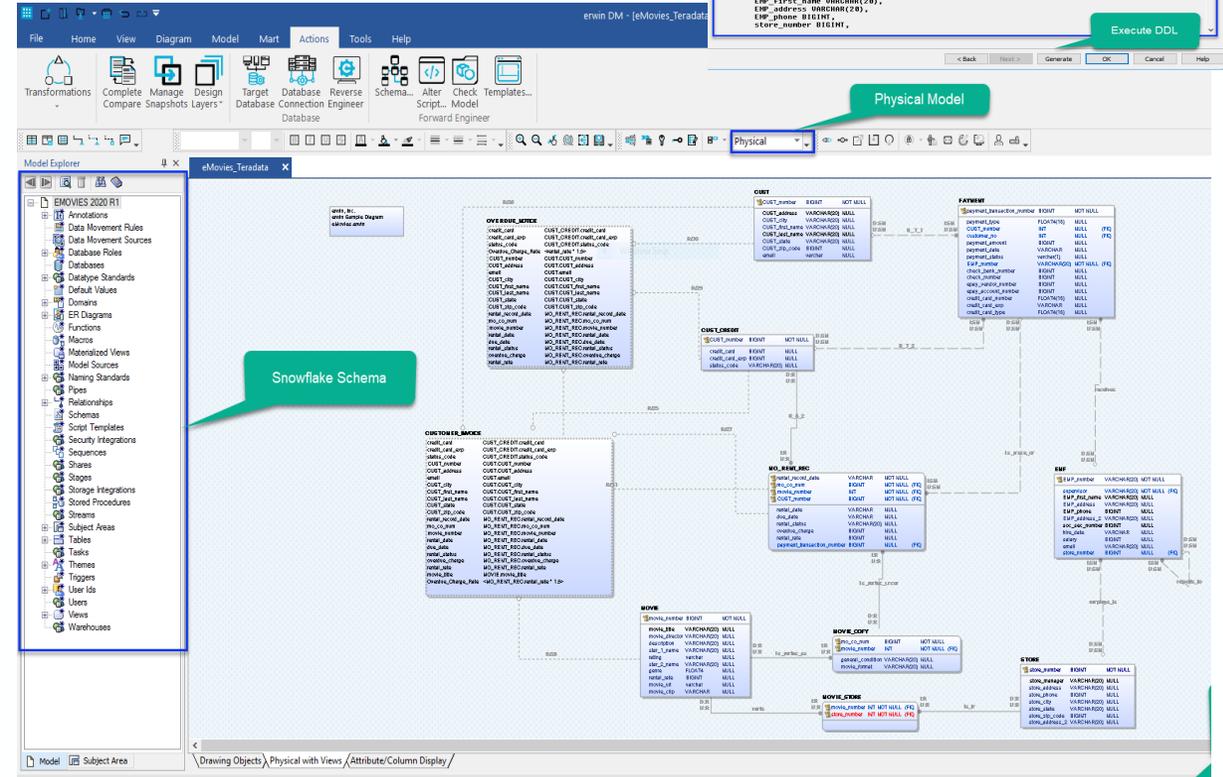
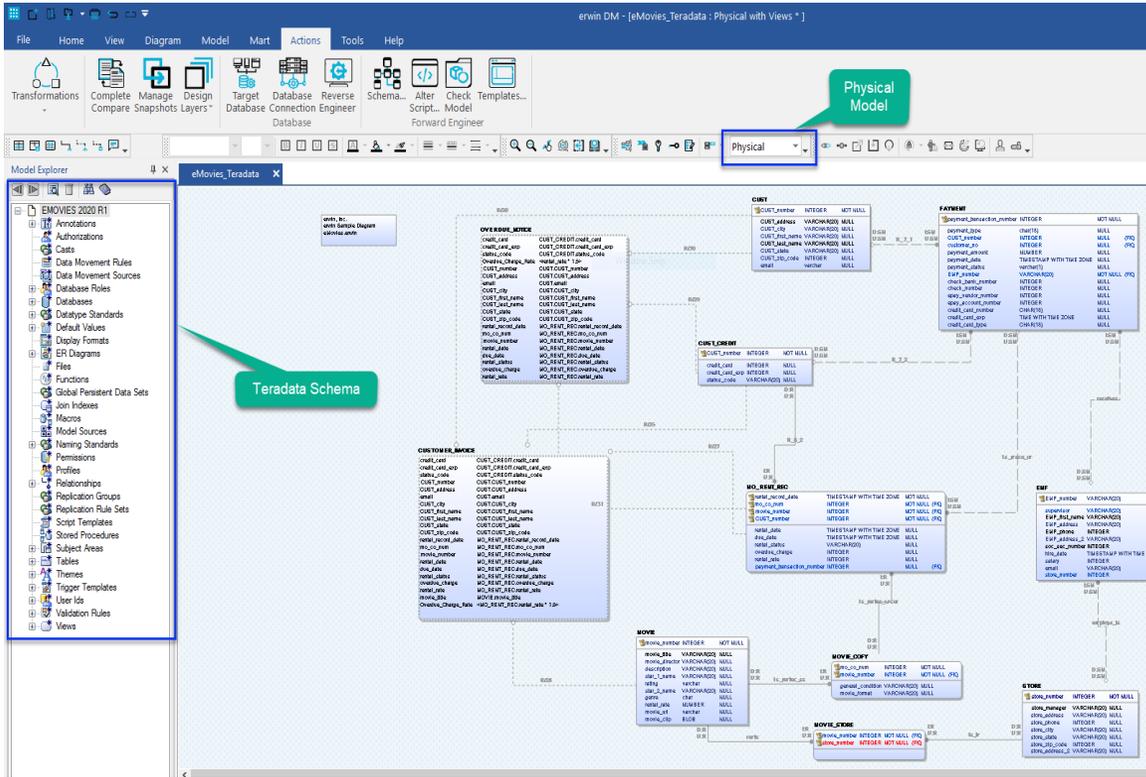
Re-Align Usage Models

Automate Dev/Ops

Continuous Data Governance

Modernizing Data Architecture

Migrate legacy schema to the cloud with data modeling

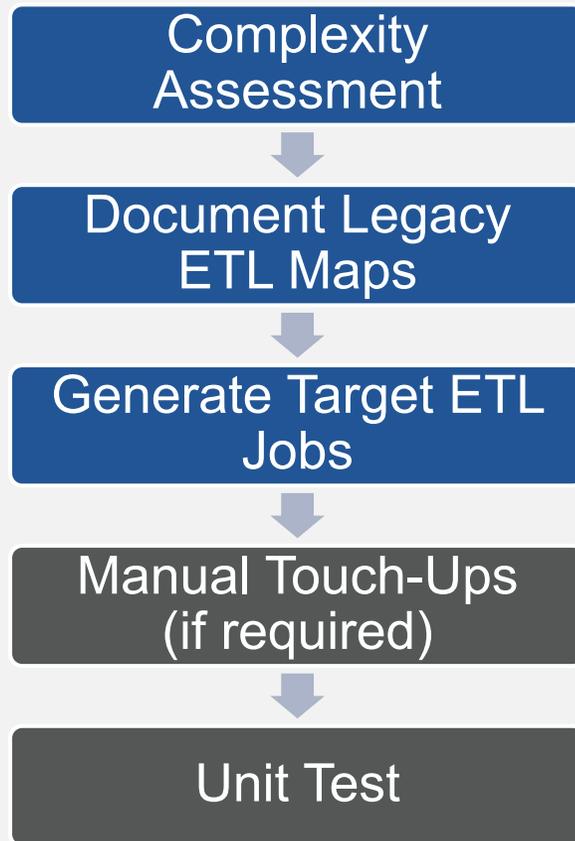


Reverse Engineer

Transform

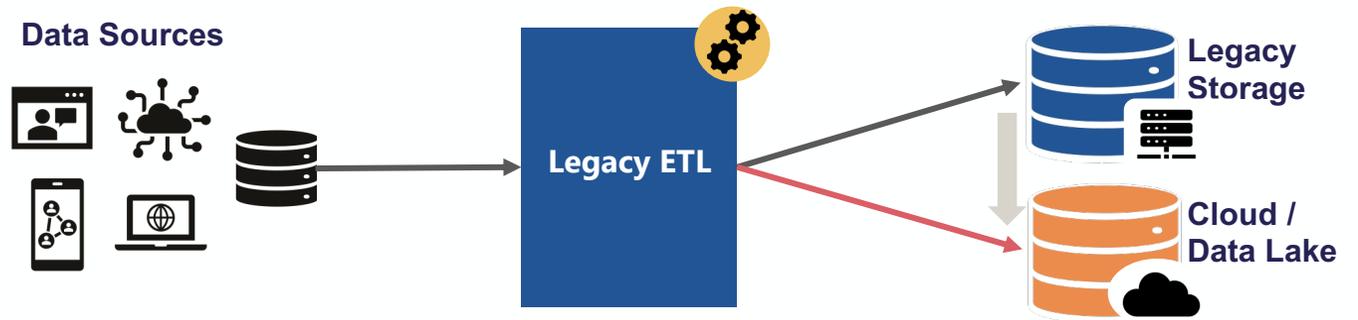
Forward Engineer

ETL Transformation



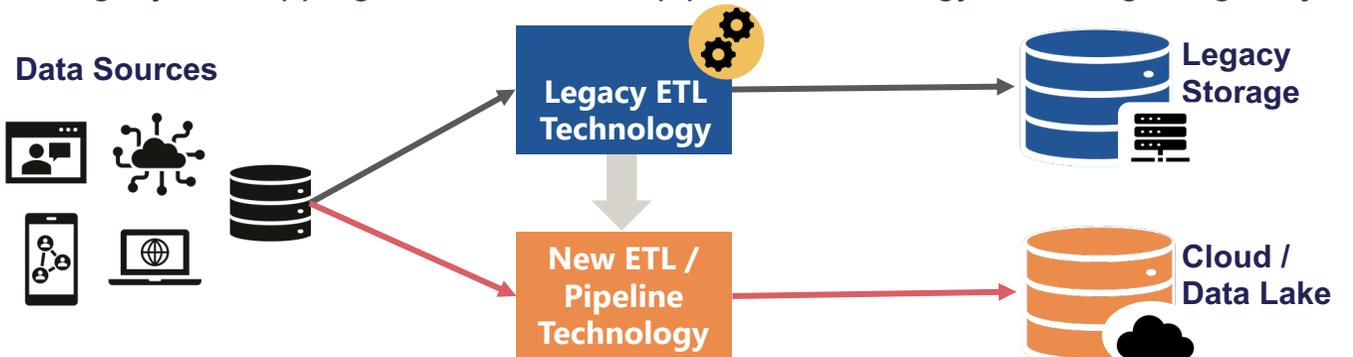
ETL Repointing & Migration

Repointing & Deploy legacy ETL jobs with new cloud environment Mappings



ETL Conversation & Migration

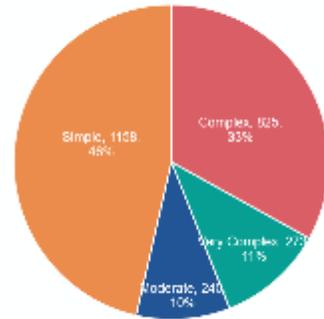
Convert legacy etl mappings to a modern etl/pipeline technology while migrating the jobs



ETL Transformation Complexity Assessment

Stratification

Complexity Distribution



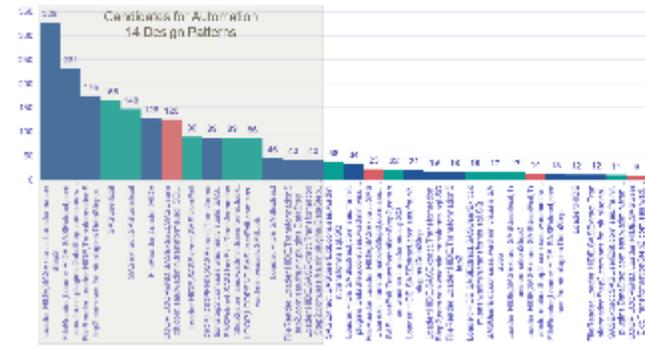
Complex Very Complex Moderate Simple

Complexity Definition

- Simple:** <=5 Components
 - example : jobs consists of one source, one target and other components
- Moderate:** <=5 Components, 1 User Written, 1 pre/postcode
 - example : mappings consists of multiple homogeneous sources (not more than 3) , multiple targets (not more than 3) or multiple transformers (not more than 3)
- Complex:** >5 and <10 Components, >=3 User Written, >=5 pre/postcode
 - example : mapping consists of multiple sources, multiple targets, multiple transformers (not more than 10), multiple mapping pipelines (not more than 5)
- Very Complex:** >10 Components, <5 User Written, <5 pre/postcode
 - example : mapping consists of multiple pipelines, complex logic or business rules, custom components, or use of sub mapping

Load Design Patterns

Candidates for Automation 14 Design Patterns



Generic Design Patterns

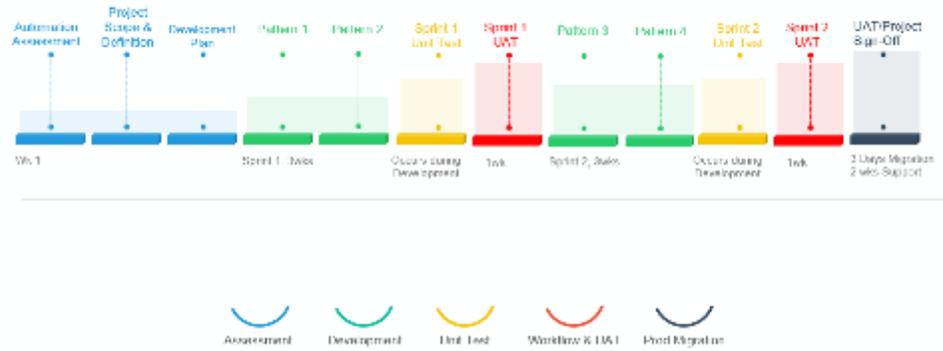
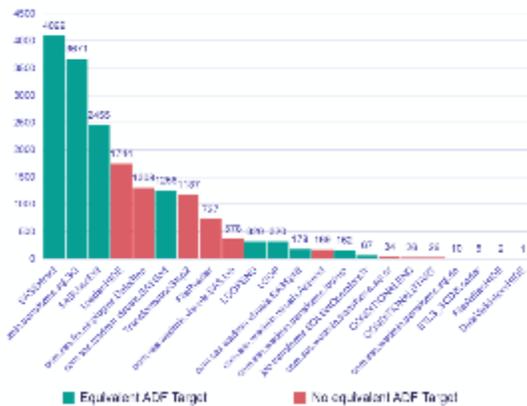
Patterns identified based on the jobs that have same distinct components are considered Generic Patterns
Candidates for automation represent 72% of the total number of jobs, although the remaining jobs will also benefit from automation, but to a lesser extent.

Component Frequency gives us an estimation of the amount of reuse to be expected across Mappings from the automated component conversion

Total Components : 22
ADF Target equivalent : 10
No equivalent ADF Target : 12
erwin must create custom components or patterns to meet the required functionality

Hidden Complexities :
SASUserExit (User Written) : 2455 – some automation possible

Component Frequency



Component Frequency

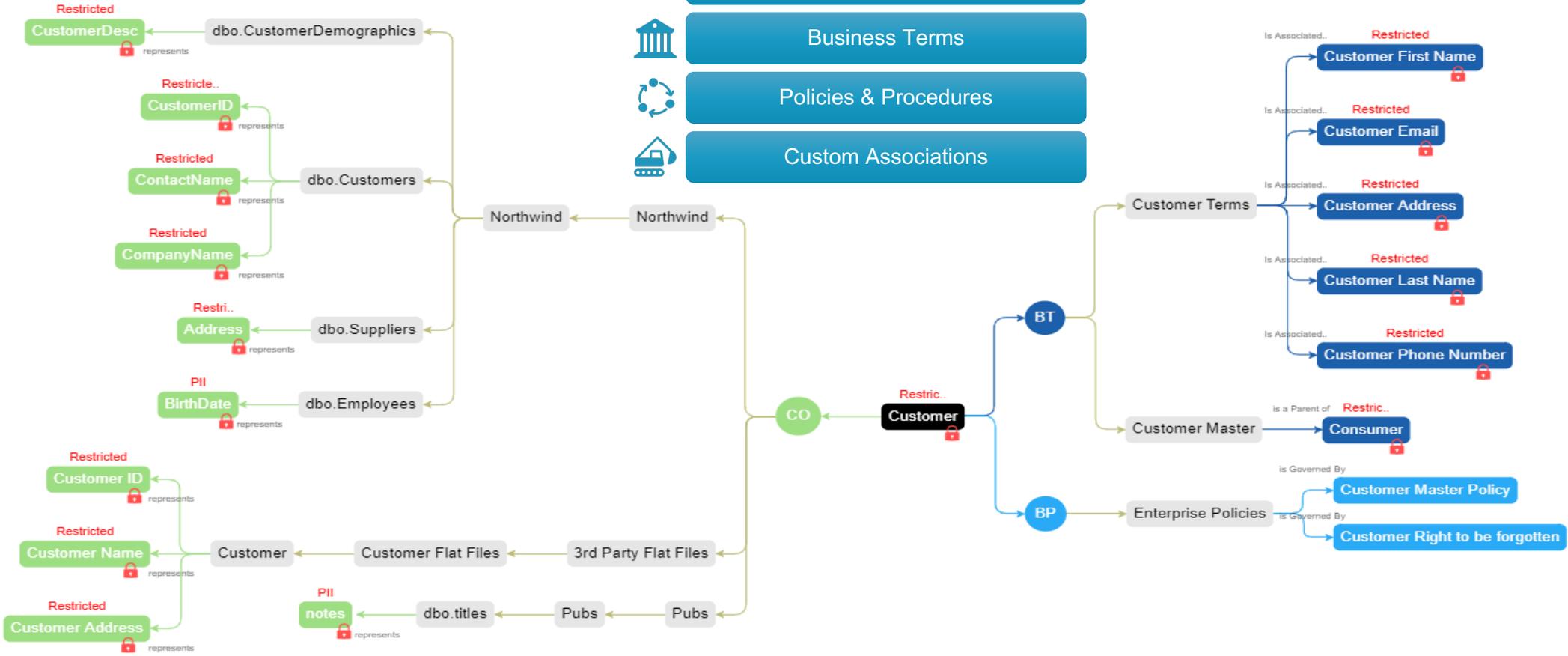
Proposed Timeline

Cloud Governance

Enabling democratization of technical assets with a Contextual Business Asset Framework

Mind Map Associations

-  Technical Assets
-  Business Terms
-  Policies & Procedures
-  Custom Associations



Cloud Governance

Automate the Discovery and Rendering of Detailed Lineage

Lineage For: ENECO_Snowflake → Snowflake-PRD → dbo.Customers → CustomerID

Overview Lineage

Property	Value
System Name	ENECO_
Environment Name	Snowflak
Table Name	dbo.Cust
Column Name	Customer
Column Data Type	nchar
Column Precision	0
Column Length	5
Column Scale	0
Xpath	
Primary Key Flag	Y
Column Identity Flag	N
Column Nullable Flag	N
Natural Key Flag	N

Cloud Governance

Automate the Classification of Sensitive Data

System Name: All Environment Name: All

Sensitivity Classification - All

- HIPPA: 6
- No Classification: 28
- PII: 11
- Public: 5
- Restricted: 13
- Secret: 79

Sensitivity Classification by System

Statistics

- 1310 Total Tables
- 18 Sensitive Tables
- 5116 Total Columns
- 142 Sensitive Columns

Summary of Sensitive Columns

#	System Name	Environment Name	Table Name	Column Name	SDI Class Name	SDI Class Description
1	Northwind	Northwind	dbo.Order Details	ProductID	HIPPA	
2	Northwind	Northwind	dbo.Orders	OrderID	HIPPA	
3	Northwind	Northwind	dbo.Orders	CustomerID	Secret	Secret
4	Northwind	Northwind	dbo.Orders	EmployeeID	Secret	Secret
5	Northwind	Northwind	dbo.Orders	OrderDate	Secret	Secret
6	Northwind	Northwind	dbo.Orders	RequiredDate	Secret	Secret
7	Northwind	Northwind	dbo.Orders	ShippedDate	Secret	Secret
8	Northwind	Northwind	dbo.Orders	ShipVia	Secret	Secret
9	Northwind	Northwind	dbo.Orders	Freight	Secret	Secret
10	Northwind	Northwind	dbo.Orders	ShipName	Secret	Secret
11	Northwind	Northwind	dbo.Orders	ShipAddress	Secret	Secret
12	Northwind	Northwind	dbo.Orders	ShipCity	Secret	Secret
13	Northwind	Northwind	dbo.Orders	ShipRegion	Secret	Secret
14	Northwind	Northwind	dbo.Orders	ShipPostalCode	Secret	Secret

**An Insights-
Driven Data
Architecture...**

**With
Intelligence,
Automation
and
Governance at
the Core**

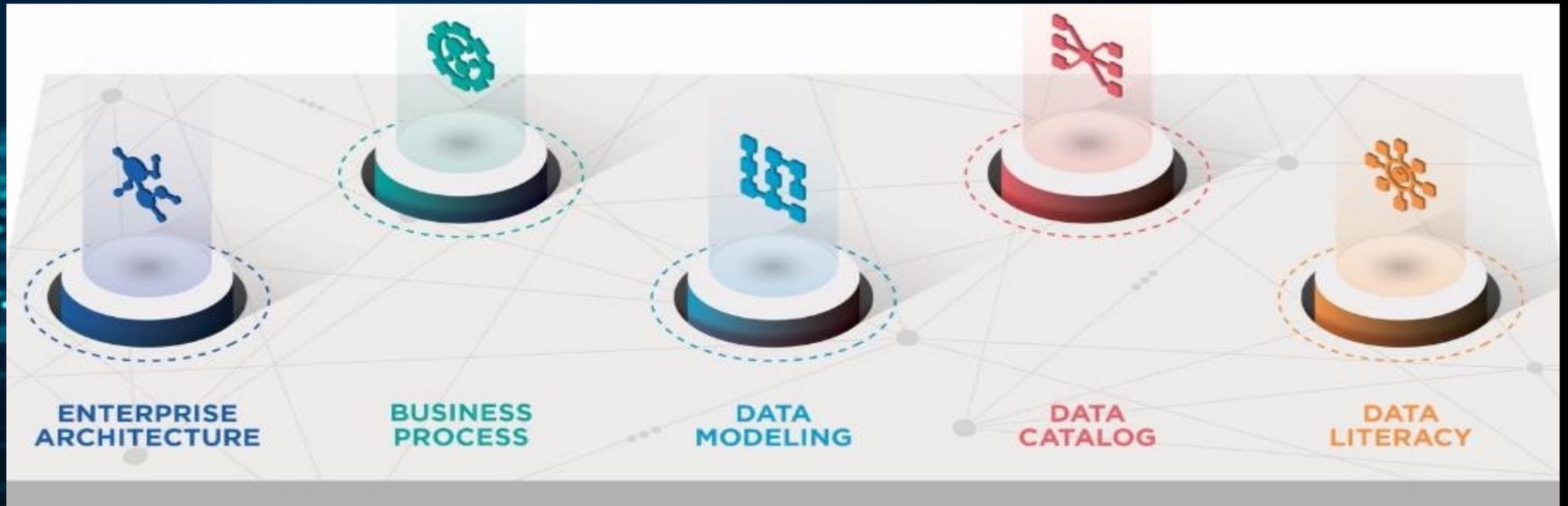
**That is as
Adaptable as
The Business
it Serves**

Questions?



EDM Webinar 

The erwin Data Intelligence Platform



Optimizing Data
Value Chains

Governing Data
Ecosystems

Automating Data
Platforms

Democratizing
Data Capability

Visit us at erwin.com

erwin[®]
by Quest

FOR MORE INFORMATION:

Danny Sandwell

Director of Product Marketing, Erwin

info@erwin.com

erwin.com

