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)
- 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
<table>
<thead>
<tr>
<th>Why Modernize? The Enterprise Data Dilemma</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% of organizations integrating at least six different types of data across 10 different data management technologies.</td>
</tr>
<tr>
<td>85% of time spent on data prep and only 15% in analysis hampers data-driven digital transformation.</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Time To Value</th>
<th>Conversion Accuracy</th>
<th>Cost Containment</th>
</tr>
</thead>
</table>

Data Governance and Intelligence

<table>
<thead>
<tr>
<th>Migration Transparency</th>
<th>Documenting cutting edge technologies</th>
<th>Data democratization enablers</th>
</tr>
</thead>
</table>
### Utilize Your Enterprise Architecture to Plan The Move

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Drivers</td>
<td>Testing</td>
<td>Cost</td>
</tr>
<tr>
<td>Risk</td>
<td>Support</td>
<td>Scope</td>
</tr>
<tr>
<td>Benefit</td>
<td>Monitoring</td>
<td>Plan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current State</td>
<td>Future State</td>
<td>Stakeholders</td>
</tr>
<tr>
<td>Complexity</td>
<td>Technologies</td>
<td>Approach</td>
</tr>
<tr>
<td>Impacts</td>
<td>Tools</td>
<td>Priorities</td>
</tr>
<tr>
<td>Requirements</td>
<td>Architecture</td>
<td>Roadmap</td>
</tr>
<tr>
<td>Security</td>
<td>Vendors</td>
<td>Capabilities</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drivers</th>
<th>Stakeholders</th>
<th>ETL Pipeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Meta data</td>
<td>Security</td>
</tr>
<tr>
<td>Scope</td>
<td>Testing</td>
<td>Monitoring</td>
</tr>
<tr>
<td>Priorities</td>
<td>Support</td>
<td>Governance</td>
</tr>
<tr>
<td>Resources</td>
<td>Future State</td>
<td>Technologies</td>
</tr>
</tbody>
</table>

**Diagram:**
- **Schema**
- **Data**
- **ETL Pipeline**
- **Meta data**
- **Apps**
Data Intelligence Enables Automation and Governance

- Harvest
- Curate
- Govern
- Activate
- Socialize
# Data Intelligence: Get More Utility From Your Metadata

<table>
<thead>
<tr>
<th>Auto Document</th>
<th>Auto Curate</th>
<th>Auto Render</th>
<th>Auto Generate</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Data Sources</td>
<td>✓ Technical Asset Associations</td>
<td>✓ End-to-End Lineage</td>
<td>✓ Data Pipelines</td>
</tr>
<tr>
<td>✓ Data Models</td>
<td>✓ Business Asset Associations</td>
<td>✓ Impact Analysis</td>
<td>✓ Data Workloads</td>
</tr>
<tr>
<td>✓ Data Movement Processes</td>
<td>✓ Sensitive Data Classification</td>
<td>✓ Mind Map Graph Views</td>
<td>✓ Data Movement Code</td>
</tr>
<tr>
<td>✓ Data Consumption</td>
<td></td>
<td>✓ Focused Dashboards</td>
<td>✓ Platform Orchestration</td>
</tr>
</tbody>
</table>
# Data Mapping Documents: Activating Metadata For Maximum Utility

## Data Movement Capture
- Scan and Auto-Document Code
- Import Mappings from Delimited Files
- Manually Specify Mappings
- Import Data Model Mappings

## Abstracted Mapping Documents
- Source
- Transformation
- Target

## Mapping Exploration and Activation
- Lineage Rendering
- Impact Analysis
- Automated Code Generation
Accelerating and Assuring Architectural Modernization

Automated documentation, transformation, code generation and governance

Data Sources
- OLTP
- Enterprise Applications
- Third-Party Applications
- Web Applications
- Other

Model

Design

Generate

Architecture
- Legacy Data Curation
- Legacy Storage
- Legacy Presentation

ETL Tools

New ETL Pipeline

New Data Platform Schemas

Cloud / Data Lake

New Information Delivery Layer

Reporting

Continuous Data Governance

Convert Data Structures

Accelerate Data Migration

Transform Data Movement

Re-Align Usage Models

Automate Dev/Ops
Modernizing Data Architecture
Migrate legacy schema to the cloud with data modeling
Data Migration Accelerators for Cloud Migration

Source to Target Mapping
- Data Migration Specification
- Manual Mappings
- Automated Mappings
- Transformations

ETL Generation
- ETL Job Generation
- Open Source / Enterprise
- Corporate Standards
- Native ETL Frameworks

Data Migration
- Native Technology Execution
- Big Data Batch
- Source Type Detection
- Target Type Detection
ETL Transformation

ETL Repointing & Migration

*Repainting & 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*

- **Complexity Assessment**
- **Document Legacy ETL Maps**
- **Generate Target ETL Jobs**
- **Manual Touch-Ups (if required)**
- **Unit Test**
ETL Transformation Complexity Assessment

**Stratification**

**Component Frequency**
- Component Frequency gives an estimation of the amount of effort to be expected across mappings from the automated component conversion.
- Total Components: 22
- ADF Target Requirement: 10
- No equivalent ADF Target: 12
- erwin must create custom components or patterns to meet the required functionality.

**Load Design Patterns**

**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
Cloud Governance
Automate the Classification of Sensitive Data
An Insights-Driven Data Architecture...

With Intelligence, Automation and Governance at the Core

That is as Adaptable as The Business it Serves
Questions?
The erwin Data Intelligence Platform

- Optimizing Data Value Chains
- Governing Data Ecosystems
- Automating Data Platforms
- Democratizing Data Capability

Visit us at erwin.com
FOR MORE INFORMATION:

Danny Sandwell
Director of Product Marketing, Erwin
info@erwin.com
erwin.com