

Going Live with Enterprise Solutions



Metadata Management is Key to Data Governance Initiatives

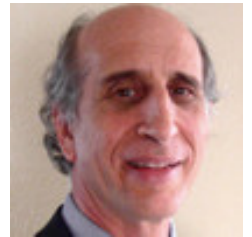
Irene Polikoff and Jack Spivak, TopQuadrant Inc.

April 27, 2017



Capabilities for Effective Metadata Management

- Why is metadata important, and what is needed to manage it?
- What can TopBraid EDG's Metadata Management Capabilities do for You?
- DEMO: EDG Metadata Management in Action
- Quick Poll
- Questions and Answers



Jack Spivak



Irene Polikoff

TOPQUADRANT COMPANY

FOUNDATION

- TopQuadrant was founded in 2001
- Strong commitment to standards-based approaches to data semantics

MISSION

- Empower people and drive results — by making enterprise information meaningful

FOCUS

- Provide comprehensive data governance solutions

ENTERPRISE CUSTOMERS



U.S. AIR FORCE



PEARSON

servicenow



THOMSON REUTERS™



VIDAL GROUP

syngenta®

AstraZeneca 



amdocs

EURIWARE
Capgemini
CONSULTING. TECHNOLOGY. OUTSOURCING

Let's talk about metadata ...

- What is it and why is it important?
- Why is it challenging to manage?
- How does a semantic standards-based approach help?
- TopBraid EDG Semantic Metadata Management in Action

Data about Data

Allows understanding of its identity and characteristics



- Name
- Origin
- Created
- Modified
- Definition
- Data type
- Source
- Valid values
- Security classification
- Security policies
- Lifecycle
- Retention policies

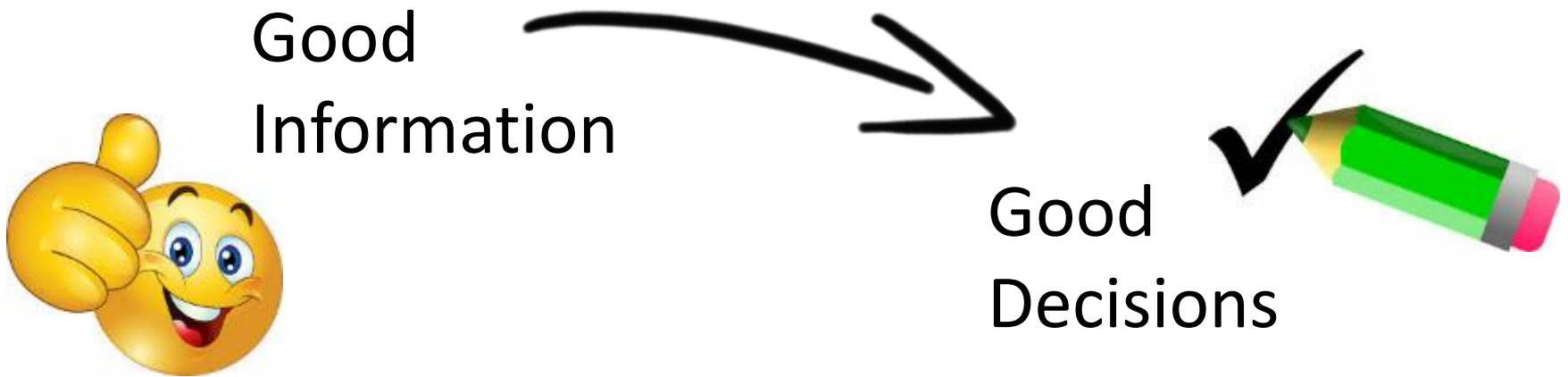
Context: business, technical, and operational

- Databases, tables, fields, procedures, functions, etc
- Software applications, scripts, procedures, functions
- Information systems, integrations, networks, hardware
- Business processes, roles, capabilities, requirements

(DATAVERSITY report: "Emerging Trends in Metadata Management: 2017 Report on the Top Business & Technical Drivers for Metadata by Donna Burbank and Charles Roe)

Good Information ...

- Correct, Reliable
- Timely
- Useful, Relevant



Meaning - allows us to make the good decisions with good information

*Good
Decisions*

+

*“Meaningful”
Information*



*Good
Outcomes*

- Intended
- Valuable
- Effective



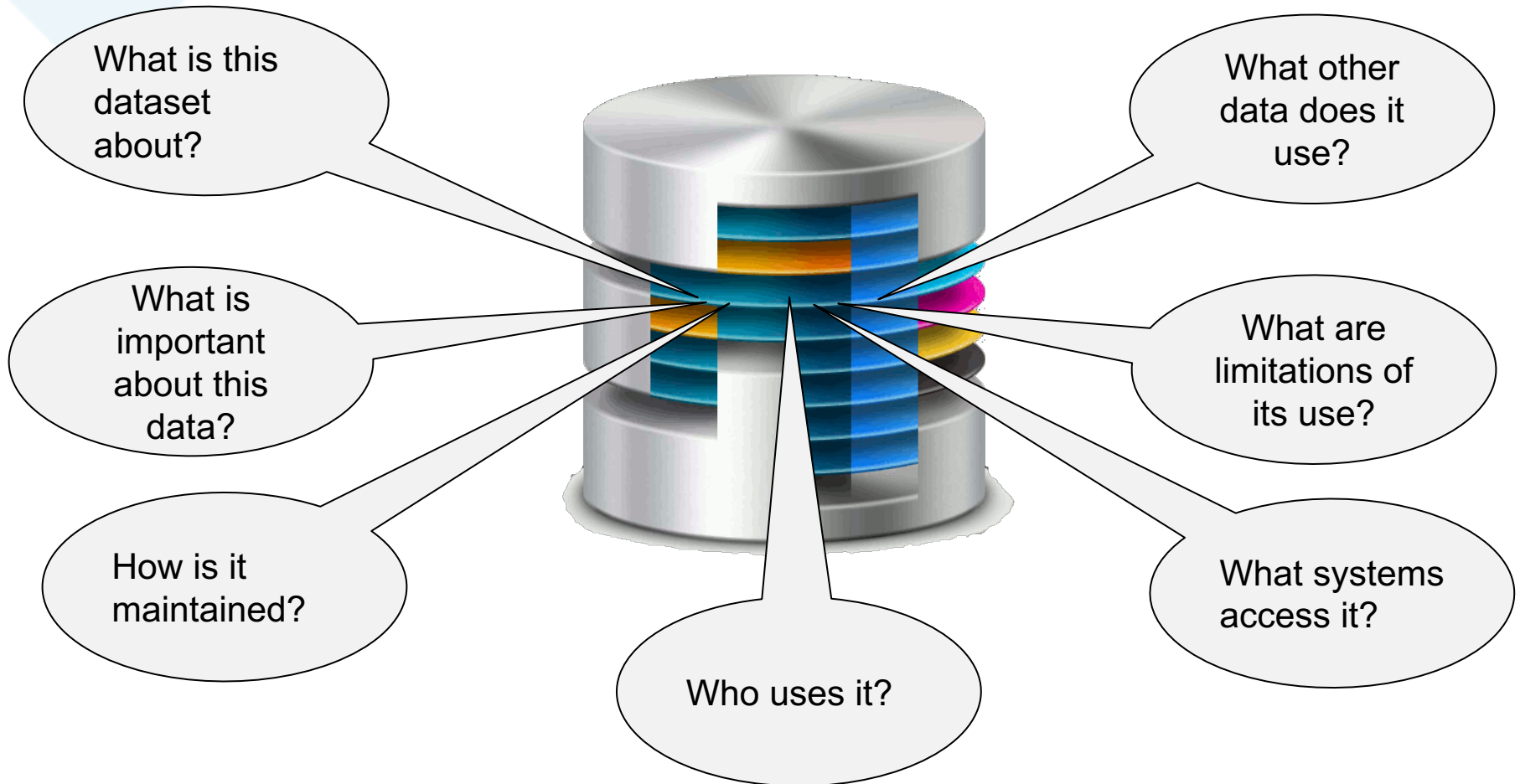
I hear thunder and
see dark clouds ...



Stay where you are and
raise your hands ...

Importance of Metadata?

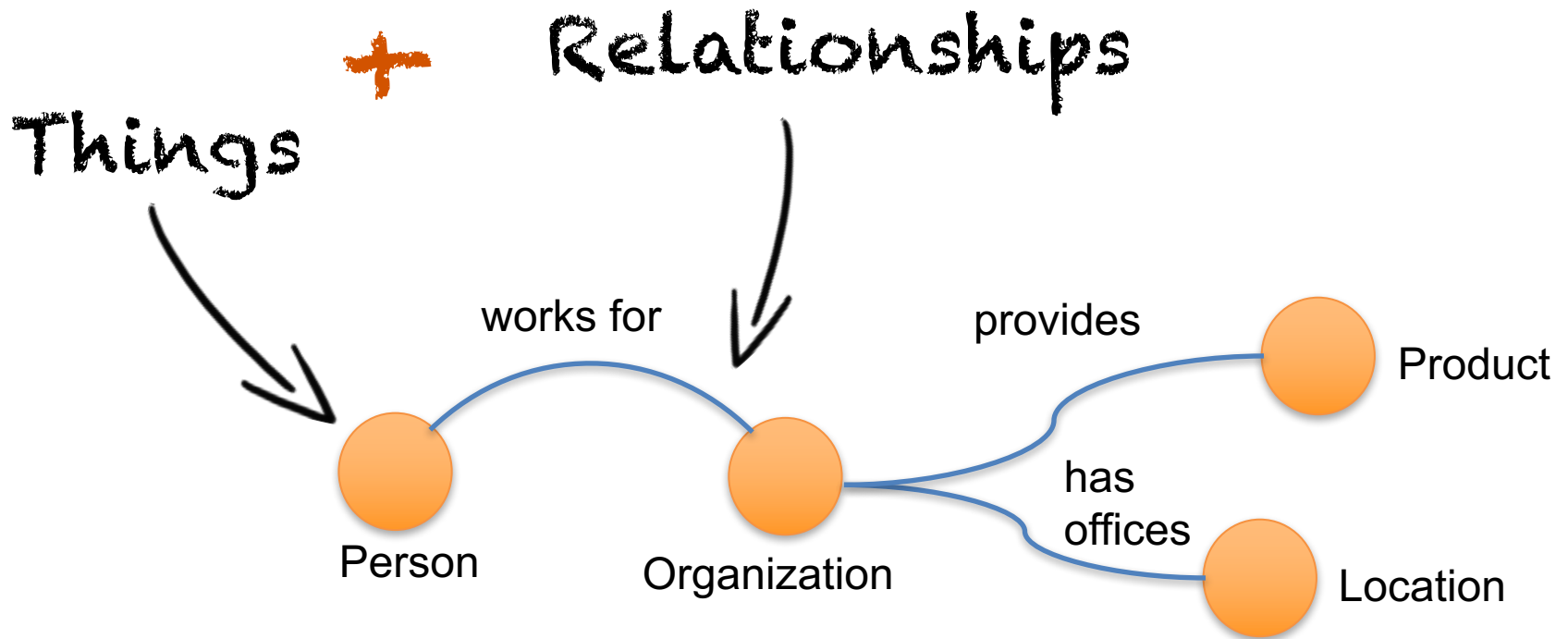
Metadata captures, communicates and clarifies the meaning of data
-- *answers key questions for better decisions*



Confusion, wasted resources and errors result when users can't understand the relation between data and business goals



Semantics =



Metadata can be of two types:



Descriptive Metadata – Values state or explain something about the data. Typically contain string, boolean or date type data.



Relational Metadata – Values connect or “point” to other objects that participate together in a process or analysis or job. This relationship makes a “model” of the actual uses, flow and impacts of data in the enterprise.

W3C* Semantic Standards

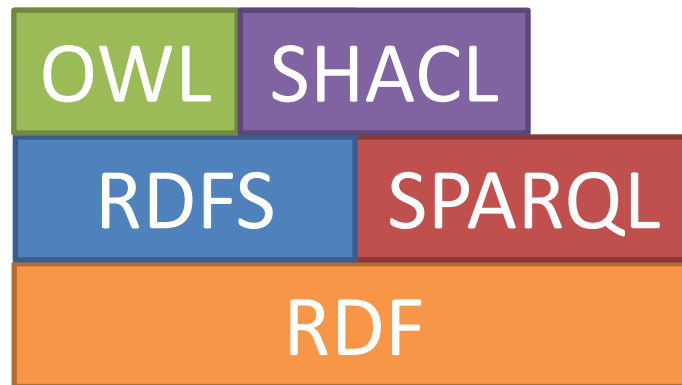
*W3C = World Wide Web Consortium
led by Tim Berners-Lee

Restricted Inference:
What is this thing?

Rules:

Is that term used correctly?
What do you need to know?
You can't say that here!

Vocabulary:
Shared terms can
we use



Query:
What did you say?

Statements:
Saying things

- Understanding
- Interoperability & Reuse
- Semantic Alignment
- Impact and Risk Analysis
- Analytics and Decision-making
- Compliance & Reporting
- Search Enhancement



Metadata Challenges

Its never easy being the
“glue” that keeps the
connections going





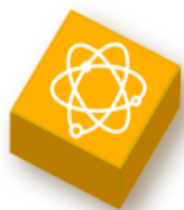
Business
Glossaries



Reference Data
Management



Basic Metadata
Management



Advanced Metadata
Management



Vocabulary
Management



Example:
Fictitious TopBank Corp

Federal Reserve Compliance
Reporting.



Form



Schedule



Data Item

Use EDG to:

- ✓ Manage Metadata at the Collection and Individual Level
- ✓ Follow Linking Metadata
- ✓ Create and Connect New Metadata
 - Descriptions
 - RACI
 - Security Classes
 - Policy References
 - Glossary w Links



Data Assets



Technical Assets



Enterprise Assets



Participant Poll

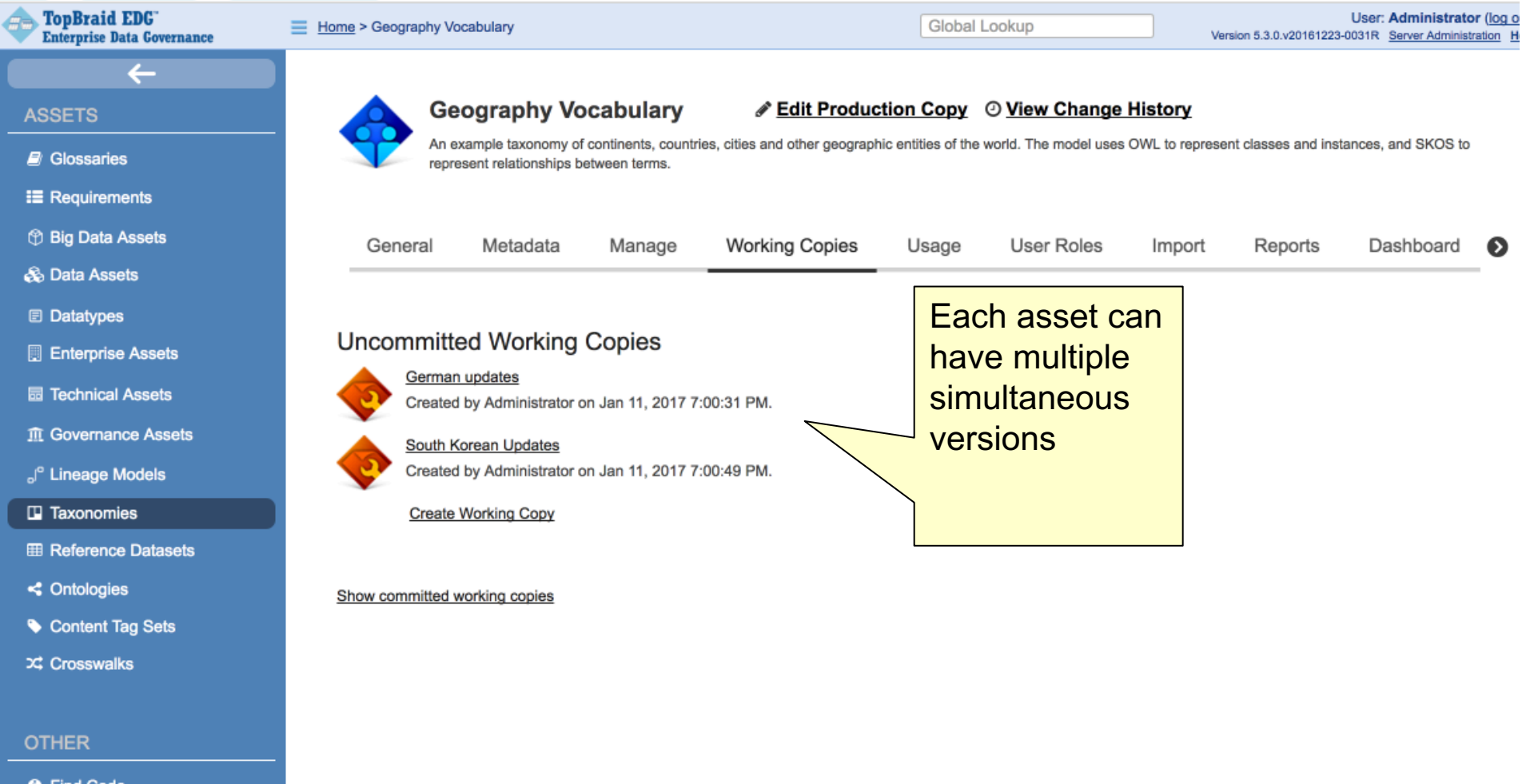
Demo: Metadata Management in TopBraid EDG

Participant Poll

- Assign users different roles for using the system at various levels of authority per asset collection:
 - Manager
 - Editor
 - Viewer



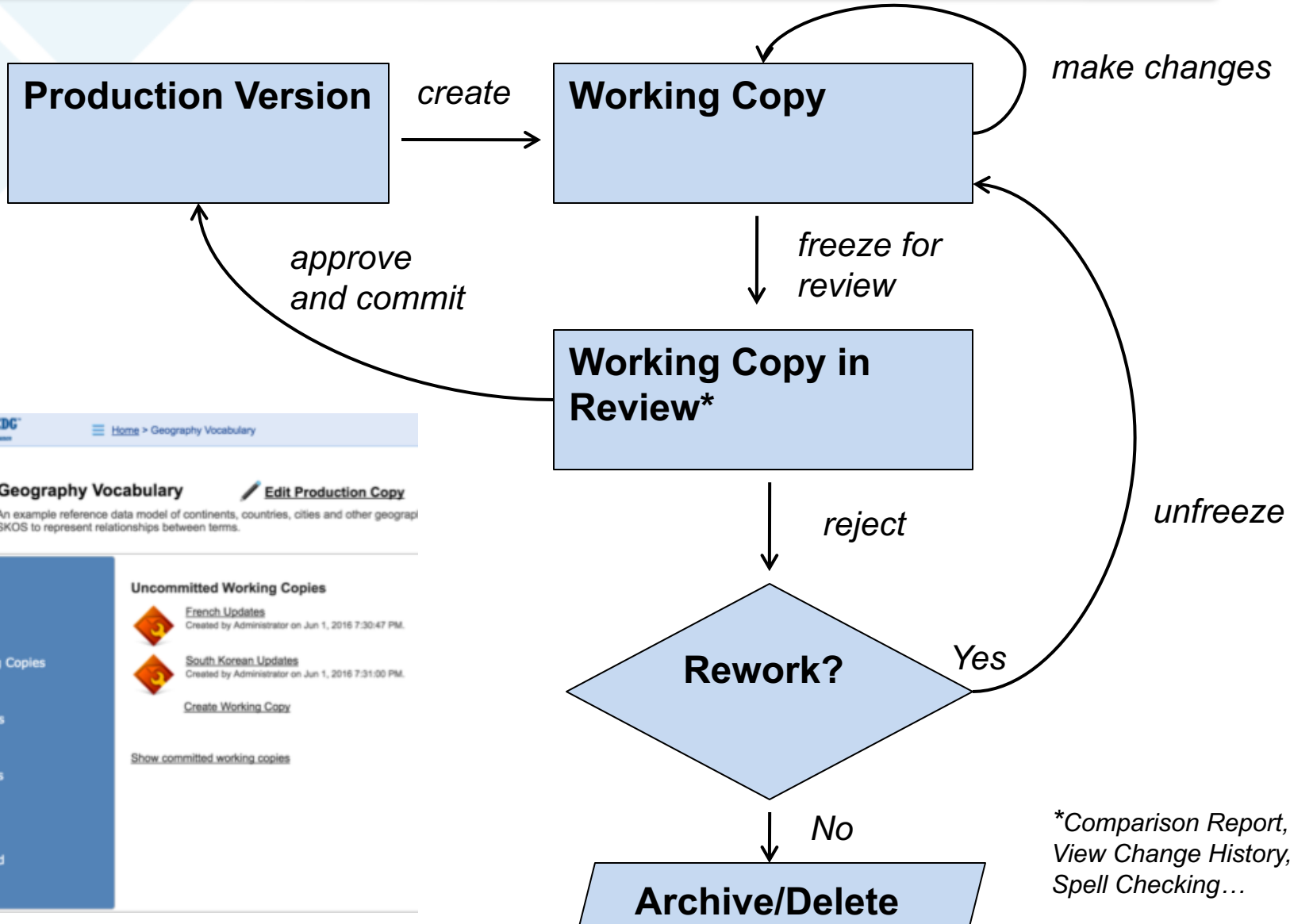
- These users are typically data stewards and curators for different asset collections
- System-wide read-only role is also available to support search and/or comment only use cases



The screenshot shows the TopBraid EDG interface. The top navigation bar includes the logo, 'Home > Geography Vocabulary', a search box, and user information. The left sidebar lists various asset types, with 'Taxonomies' selected. The main content area displays the 'Geography Vocabulary' page, which includes a description and a set of tabs: General, Metadata, Manage, Working Copies (selected), Usage, User Roles, Import, Reports, and Dashboard. Under the 'Working Copies' tab, there is a section for 'Uncommitted Working Copies' listing two entries: 'German updates' and 'South Korean Updates', each with a creation timestamp. A 'Create Working Copy' link is also present. A yellow callout box with a pointer to the list states: 'Each asset can have multiple simultaneous versions'.

Each item has a separate audit trail from multiple users

Built-in Workflow Support



*Comparison Report, View Change History, Spell Checking...

TopBraid EDC™
Enterprise Data Governance

Home > Taxonomies > Geography Vocabulary > Editor

User: Administrator (log out)
Version 5.3.0.v20170417-1101B Server Administration Help

Concept Hierarchy

Germany

- Geography
 - Africa
 - Antarctica
 - Asia
 - Atlantic Ocean
 - Europe
 - Albania
 - Andorra
 - Austria
 - Belarus
 - Belgium
 - Bosnia-Herzegovina
 - Bulgaria
 - Croatia
 - Czech Republic
 - Denmark
 - Estonia
 - Finland
 - France
 - Germany**
 - Gibraltar
 - Greece
 - Greenland
 - Hungary
 - Iceland
 - Ireland
 - Italy
 - Latvia
 - Liechtenstein
 - Lithuania
 - Luxembourg

Germany (Country)

Labels and Description

preferred label:

- Alemania (es)
- Allemagne (fr)
- Deutschland (de)
- Duitsland (nl)
- Germany (en)
- Niemcy (pl)
- Германия (ru)

type: [Country](#) →

Label Objects (SKOS-XL)

alternative label (object):

- [Allied-occupied Germany](#) →
- [East Germany](#) →
- [Federal Republic Of Germany](#) →
- [German Democratic Republic](#) →
- [German Empire](#) →
- [German Reich](#) →
- [Nazi Germany](#) →
- [Third Reich](#) →
- [Weimar Republic](#) →
- [West Germany](#) →

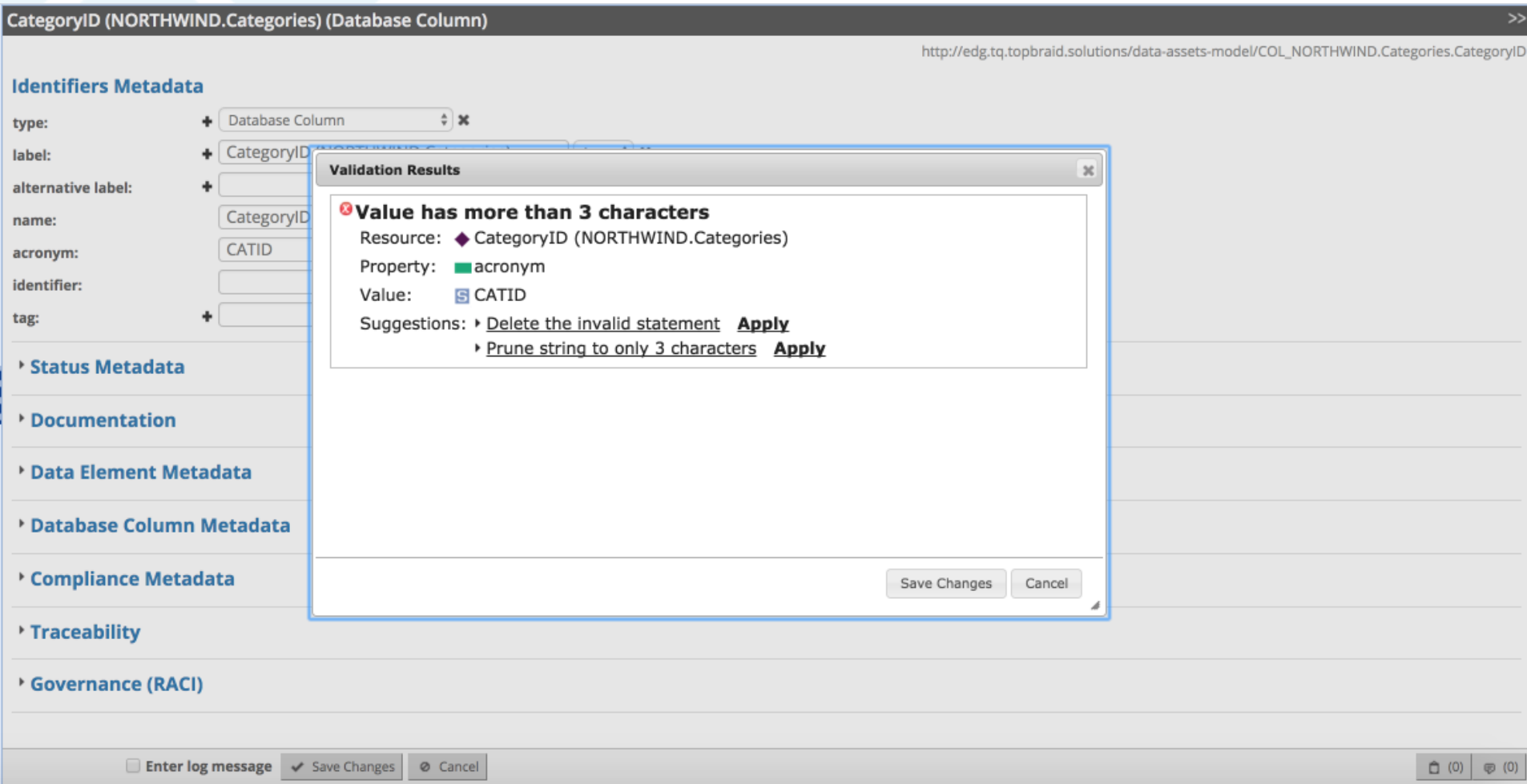
Standard Relationships

has broader: [Europe](#) →

has related: [European Union](#) →

Custom Properties

Show History

The screenshot shows a web-based interface for managing metadata. The main window is titled "CategoryID (NORTHWIND.Categories) (Database Column)". On the left, there is a sidebar with various metadata categories: Identifiers Metadata, Status Metadata, Documentation, Data Element Metadata, Database Column Metadata, Compliance Metadata, Traceability, and Governance (RACI). The "Identifiers Metadata" section is active, showing fields for type, label, alternative label, name, acronym, identifier, and tag. The "acronym" field contains the value "CATID". A "Validation Results" dialog box is open in the center, displaying an error message: "Value has more than 3 characters". The dialog provides details: Resource: CategoryID (NORTHWIND.Categories), Property: acronym, and Value: CATID. It offers two suggestions: "Delete the invalid statement" and "Prune string to only 3 characters", each with an "Apply" button. At the bottom of the dialog are "Save Changes" and "Cancel" buttons. The main interface also has a footer with "Enter log message", "Save Changes", and "Cancel" buttons, and a status bar on the right showing "(0)" for two different icons.

For information governed by EDG e.g., reference data, glossary terms, etc.

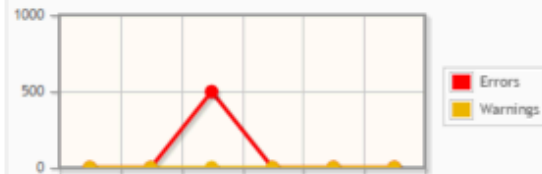
Completeness and Validity

✔ Data Completeness: 100%

0 data entries out of 0 are incomplete, i.e. they lack required properties.

⚠ Data Validity over Time

6 reports shown.
Most recent report from Jan 8, 2016 7:22:43 AM:
1 Errors, 0 Warnings



Information about the Reference Dataset

⚠ Completeness of Descriptive Metadata: 75%

Some recommended descriptive metadata is missing

- Description present
- Identifier present
- Official name present
- Purpose not present

⚠ RACI Completeness: 0%

Some recommended RACI roles are missing

- Accountability not assigned
- Responsibility not assigned
- SME (consulted) not assigned

✔ External Dataset Profiling

Airports is not an external dataset. Therefore, no Subscription properties need to be filled in.

✔ Semantic Analysis of Airport: 100%

The main entity (class) of this dataset has 5 properties. 0 of these properties lack label or description.

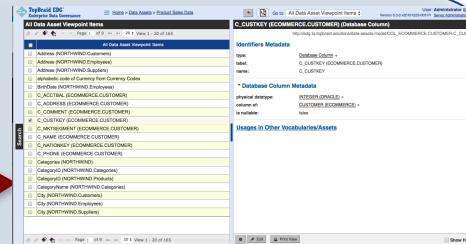


Import formats

- Spreadsheets (xls, csv, tsv)
- RDF
- XML
- UML
- RDBMs and SQL DDL
- System specific outputs e.g., Collibra, IBM Bus Glossary, MultiTes

TopBraid EDG

govern information assets



Export formats

- HTML
- tab-delimited files
- RDF
- Avro, JSON
- Sharepoint, etc.

And also export to TopBraid EDG Explorer module

TBL Server

SPARQL

Rules

Data transforms

Import & Export

EDG Repository

Data Integration services

Web Services

direct RDF/SPARQL access

JMS

Web Services

Real Time Access

3rd party systems

Data Warehouse

Portals

In-House Apps

Data Stores

Import/Export formats

- RDF (Turtle, RDF/XML, JSON-LD)
- relational databases
- XML, XMLS
- RSS
- Excel spreadsheets
- delimited text (.csv, tsv)
- Web pages (RDFa, Microformats)

TopBraid EDG Meta-model



EDG Data Assets Collections



EDG Technical Assets Collections



EDG Enterprise Assets Collections



Enterprise Data Landscape:

Linking Business, Technical, Operational Contexts

TopBraid EDG Data Governance Packages



Business Glossaries

Support for
Business Glossaries
Ontologies



Reference Data Management

Support for
Reference Datasets
Ontologies
Crosswalks



Vocabulary Management

Support for
SKOS Taxonomies
Ontologies
Crosswalks
Content Corpora



Basic Metadata Management

Support for
Data Assets
Datatypes
Ontologies



Advanced Metadata Management

Support for
Data Assets
Big Data Assets
Datatypes
Lineage Models
Technical Assets
Enterprise Assets
Requirements
Ontologies

For details, see:

<http://www.topquadrant.com/products/topbraid-edg-gov-packs/>

- *Core flexibility and extensibility* – add user defined models, assets and properties as needed (model-driven)
- *Models: pre-built and user defined* – support multiple types of governance assets
- *Connections:* – can be made between any types of assets
- *Flexible Connections Enable:*
 - People (UI) and software (APIs/web services) to view, follow and query the connections to answer core questions, e.g. *“Where did this come from?”*
 - complete data governance vs. siloed data governance, i.e *“reference-ability”*



... Questions?

- *Core flexibility and extensibility* – add user defined models, assets and properties as needed (model-driven)
- *Models: pre-built and user defined* – support multiple types of governance assets
- *Connections:* – can be made between any types of assets
- *Flexible Connections Enable:*
 - People (UI) and software (APIs/web services) to view, follow and query the connections to answer core questions, e.g. *“Where did this come from?”*
 - complete data governance vs. siloed data governance, i.e *“reference-ability”*

TopBraid EDG Data Governance Packages



Business Glossaries

Support for
Business Glossaries
Ontologies



Reference Data Management

Support for
Reference Datasets
Ontologies
Crosswalks



Vocabulary Management

Support for
SKOS Taxonomies
Ontologies
Crosswalks
Content Corpora



Basic Metadata Management

Support for
Data Assets
Datatypes
Ontologies



Advanced Metadata Management

Support for
Data Assets
Big Data Assets
Datatypes
Lineage Models
Technical Assets
Enterprise Assets
Requirements
Ontologies

For details, see:

<http://www.topquadrant.com/products/topbraid-edg-gov-packs/>

To Learn More ...

Contact us: at info@topquadrant.com to:

- Discuss metadata management solutions
- Request a more targeted demo of TopBraid EDG
- Ask for a free EDG evaluation account

EDG Metadata Management Product Info:

- <http://www.topquadrant.com/products/topbraid-edg/>
- <http://www.topquadrant.com/products/topbraid-edg-gov-packs/>

Other EDG demos/webinar recordings:

- <http://www.topquadrant.com/knowledgeassets/videos/#edgoverviewdemo>
Webinar: Data Governance for the Connected Enterprise:
TopBraid EDG in Action

Upcoming Webinar...

May 18

GDPR in EDG

**Semantic Data Governance
for GDPR Compliance**

LIVE WEBINAR

Tuesday, May 18th @ 11:30 a.m. EST

Are you ready for GDPR? If you're like most organizations, the answer is probably no. But with the compliance deadline nearing, the pressure to comply is high. **Register now!**



topquadrant.com/gdpr-compliance