

Going Live with Enterprise Solutions



Data Governance for the Connected Enterprise

Irene Polikoff and Jack Spivak, TopQuadrant Inc.

November 3, 2016



Today's Program

- I. *Capabilities for Effective Data Governance:
Why TopBraid EDG?*
- II. *DEMO: EDG Data Governance in Action*
- III. *Quick Poll*
- IV. *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



servicenow



THOMSON REUTERS™



VIDAL GROUP

syngenta®

AstraZeneca 



amdocs

EURIWARE
Capgemini
CONSULTING. TECHNOLOGY. OUTSOURCING

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

- Diverse Data Storage

- Relational Data
- Data Warehouses
- Datasets
- NoSQL Databases

- Diverse Data Models

- Database Definition Files
- Held in programs
- Held in spreadsheets
- Silo-ed data models

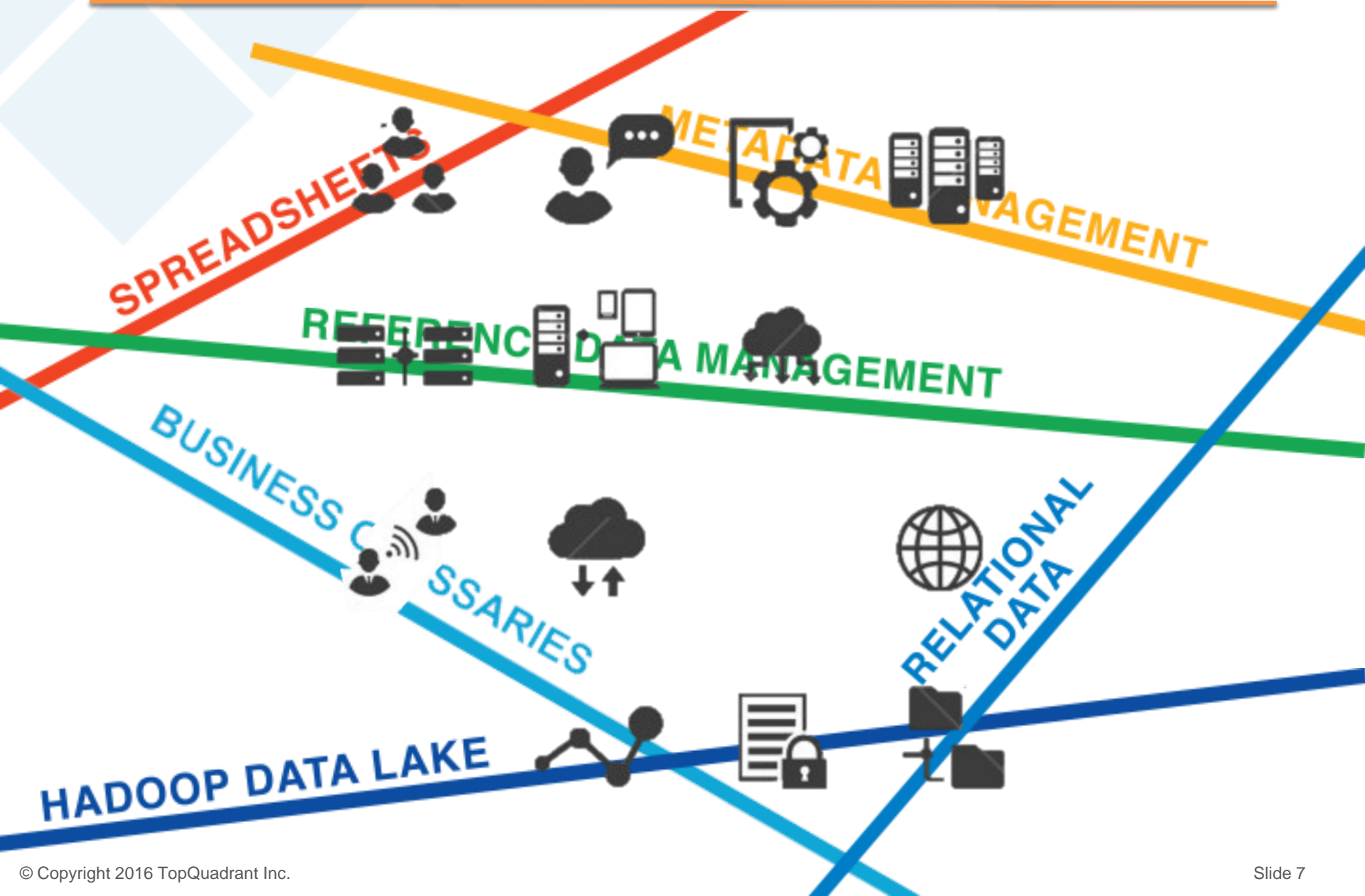
- Diverse Governance Needs

- Quality
- Policy
- Metrics
- Management

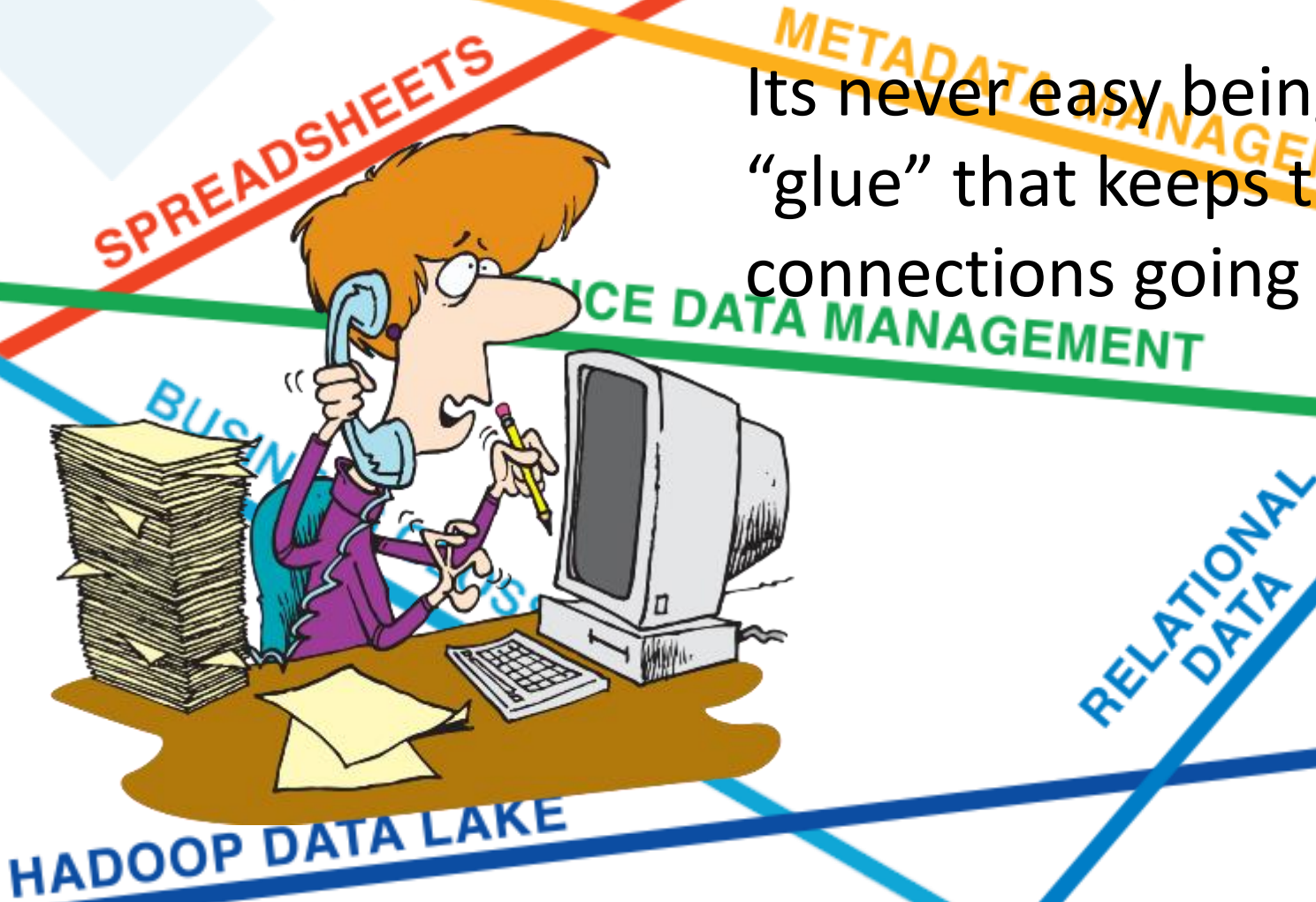
- Diverse Executables
(also worthy of governance)

- Business Rules
- Big Data Jobs
- Critical Algorithms
- Transforms

Diverse Data Governance Challenges



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



Standards-based, model-driven approach

Technical

- Comprehensive
 - Incremental
 - Flexible and Adaptable
- Operations Business

SEMANTIC INFORMATION MANAGEMENT





TRADITIONAL



BUSINESS **OR** TECHNICAL
METADATA



RIGID AND LIMITED
INFORMATION
STRUCTURES



VENDOR SPECIFIC
INFORMATION
MODELS

SEMANTIC



BUSINESS **AND** TECHNICAL
METADATA



RICH AND FLEXIBLE



CAPTURE ANYTHING



CHANGE AND EXTEND
WITHOUT INVOLVING IT



STANDARDS BASED



INTEROPERABLE



- **Metadata management** – connect technical metadata across different data sources, map physical to logical to conceptual models and to glossary terms
- **Business glossaries** – SKOS based and customized structures
- **Reference data management** – govern and distribute code sets (countries, currencies, etc.) to applications
- **Data lineage tracking** – trace lineage from context of use to origin across data stores
- **Big data management** – register, connect and locate dataset in the data lakes

How data governance with Topbraid EDG meets critical business needs:

Meaningful Data

- Helps business users understand and trust their data

Purposeful Analytics and Reporting

- Provides timely access to available data
- Confirms that data comes from the correct sources

Accurate Regulatory Compliance Submissions

- Proves where reported data originates
- Ensure data contains only permissible values

TopBraid EDG: Key Capabilities for Effective Data Governance

- Ability to flexibly represent a broad range of data assets including databases, tables and columns
- Ability to flexibly document connections – from any type of data asset to any other type of data asset – at any level
- Ability to support collaborative data governance

Demo

- Ability to represent a broad range of data assets
 - Add or delete metadata fields
 - Add new types of data assets
- Ability to connect any type of data asset to any other type of data asset – at any level
 - for example, connect business metadata with technical metadata as needed for lineage tracking and compliance reporting
- Ability to support collaborative data governance
 - Multi-user role based workflow and notifications
 - Change history and audit tracking
 - “working copies” for safely managing changes to glossaries, reference data, controlled vocabularies ...
- Model-driven configurability – user experience and behavior

TopBraid EDG Data Governance Packages



Business Glossaries

Support for

Business Glossaries



Reference Data Management

Support for

Reference Datasets

Ontologies

Crosswalks



Vocabulary Management

Support for

SKOS Taxonomies

Ontologies

Crosswalks



Basic Metadata Management

Support for

Data Assets

Datatypes



Advanced Metadata Management

Support for

Data Assets

Big Data Assets

Datatypes

Lineage Models

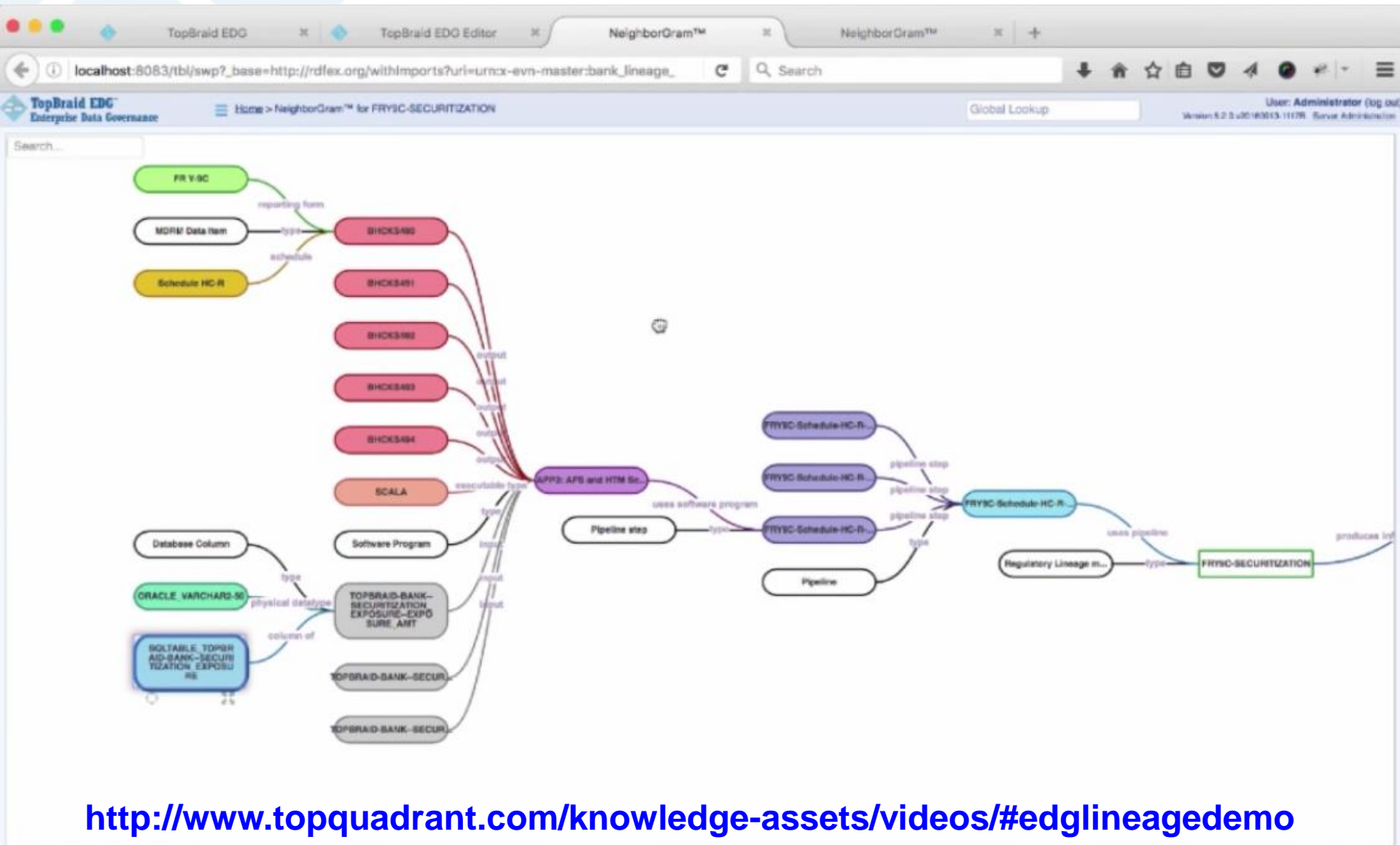
Technical Assets

Enterprise Assets

For details, see:

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

Participant Poll



<http://www.topquadrant.com/knowledge-assets/videos/#edglineagedemo>

TopBraid Enterprise Data Governance Reference Dataset Market Identifier Codes

Market Identifier Code

Market Identifier Code	Name-Institution Descr	Operating MIC	Comment	City	O/S
360T	360T	360T		FRANKFURT	Operating (O)
AATS	ASSENT ATS	AATS	ASSENT ATS IS A US	JERSEY CITY	Operating (O)
ACEX	ACE DERIVATIVES & C	ACEX	DERIVATIVES EXCHA	MUMBAI	Operating (O)
AFET	AGRICULTURAL FUTLAFET				
AIXE	AIXECUTE	AIXE			
ALDP	NYSE ALTERNEXT DAXNYS				
ALTX	ALTERNATIVE EXCHA XJSE				
ALXA	NYSE EURONEXT - ALXAMS				
ALXB	NYSE EURONEXT - ALXB				
ALXL	NYSE EURONEXT - ALXLIS				
ALXP	NYSE EURONEXT - ALXP				
AMTS	MTS NETHERLANDS XMTS				

ALXB (Market Identifier Code)

mic id: ALXB
 name-institution description: NYSE EURONEXT - ALTERNEXT BRUSSELS
 operating MIC: XBRU

TopBraid Enterprise Data Governance Home

Reference Datasets Ontologies Crosswalks Code Lookup Browse Search Tasks Organizations Software Systems

Instances of Software System

Software System

Atlas

Label: Atlas
 used by: Sales

Roles

- system business owner: Trading Platform (Organization)
- system business SME: Investment Bank CDO (Organization)
- system IT owner: Person 1 (Person)
- system technical SME: Person 2 (Person)

Usage Records

Vocabulary	Reported by	Start date	End date
Country Codes	Administrator	Jun 1, 2015	
Country Subdivision Code	Administrator	Dec 5, 2015	
Currency Codes	Administrator	Dec 5, 2015	

Compliance Reports

Vocabulary	Time stamp	Number of violations

TopBraid Enterprise Data Governance Home

Status: Under development Last updated: Oct 13, 2015 4:35:39 PM

Completeness and Validity

Data Completeness: 66%
 2 data entries out of 6 are incomplete, i.e. they lack required properties

Data Validity over Time
 7 reports shown. Most recent report from Oct 15, 2015 10:57:06 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: 33%
 Some recommended RACI roles are missing

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

<http://www.topquadrant.com/knowledge-assets/videos/#usingrdm5.0>

Status: Under development Last updated: Oct 13, 2015 4:35:39 PM

Completeness and Validity

Data Completeness: 66%

2 data entries out of 6 are incomplete, i.e. they lack required properties.



Data Validity over Time

7 reports shown.
Most recent report from Oct 15, 2015 10:57:06 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: 33%

Some recommended RACI roles are missing

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

✘ Value is not in [ex:Red, ex:Orange, ex:Green]

Resource: Traffic Light 2 (URI: http://example.org/vocabulary1/Traffic_Light_2)
Property: color (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#color>)
Value: Blue (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#Blue>)
Suggestions:

- ▶ [Delete the invalid value](#) **Apply**
- ▼ [Replace invalid value with Red](#) **Apply**
 - ✘ Blue (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#Blue>)
 - ✚ Red (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#Red>)
- ▶ [Replace invalid value with Orange](#) **Apply**
- ▶ [Replace invalid value with Green](#) **Apply**

✘ Property has more than 1 values (2)

Resource: Traffic Light 2 (URI: http://example.org/vocabulary1/Traffic_Light_2)
Property: color (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#color>)
Suggestions:

- ▶ [Delete 1 of the 2 values: Blue](#) **Apply**
- ▼ [Delete 1 of the 2 values: Orange](#) **Apply**
 - ✘ Orange (URI: <http://datashapes.org/tosh/tests/suggestions/in-001.test#Orange>)

⚠ Value has more than 10 characters

Resource: MultiLabelInstance (URI: <http://example.org/vocabulary1/MultiLabelInstance>)
Property: label (URI: <http://www.w3.org/2000/01/rdf-schema#label>)
Value: MultiLabelInstance (Datatype: string)
Suggestions:

- ▶ [Delete the invalid value](#) **Apply**
- ▼ [Prune string to only 10 characters](#) **Apply**
 - ✘ MultiLabelInstance (Datatype: string)
 - ✚ MultiLabel (Datatype: string)

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

Import formats

- Spreadsheet hierarchies (xls, csv, tsv)
- RDF
- XML
- MultiTees



TopBraid EDG

govern data assets

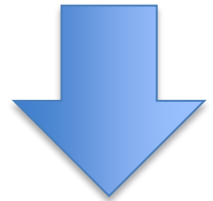


Export formats

- HTML
- tab-delimited files
- RDF
- Sharepoint



And also export to TopBraid EDG Explorer module



TBL Server

SPARQL

Rules

Data transforms

Import & Export

EDG Repository

Data Integration services

Web Services

Real Time Access

3rd party systems

Data Warehouse

Portals

In-House Apps

Data Stores

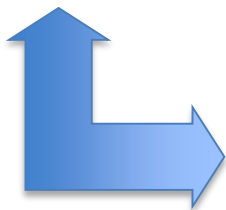
direct RDF/SPARQL access

JMS

Web Services

Import/Export formats

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





... 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: pre-built and user Defined* – 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”*

To Learn More ...

- Visit www.topquadrant.com :
 - <http://www.topquadrant.com/products/topbraid-edg/>
 - <http://www.topquadrant.com/products/topbraid-edg-gov-packs/>
- Contact us to:
 - Discuss data governance solutions
 - Request a more targeted demo of TopBraid EDG
 - Request a free evaluation account:
<http://www.topquadrant.com/products/topbraid-enterprise-data-governance/request-edg-evaluation-account/>

E-mail: edg-info@topquadrant.com

Phone: 919.300.7945