ASSET DATA MUST BE **FAIR**

Towards data driven knowledge based decision support

Jan Voskuil | Taxonic

LDAC, 17-19 June 2020

taxonic.com

WHO WE ARE

Cornelis Versteegt (PhD)
Transport and Logistics

Jan Voskuil (PhD)
Semantic Web and FAIR Evangelist

MACOMI

Founded: in 2012 with a focus on process optimization and process analytics

Tactical planning – ProRail, Port of Rotterdam

Analytics – Macomi Platform

TAXONIC

Founded: in 2012 with a strong commitment to model driven architectures

Asset Management – Schiphol

Knowledge Graph technology – TopQuadrant Partner

Disclaimer

- Taxonic has delivered and implemented an RDF-based asset data platform to Schiphol Airport
- Today, this project is only discussed in general terms
- We discuss specific features of our **productized solution**: these are generic
- No claims are made as to whether these features are implemented at Schiphol, let alone how

OUR VISION

ASSET MANAGEMENT
DONE RIGHT:

EVIDENCE BASED

>Data

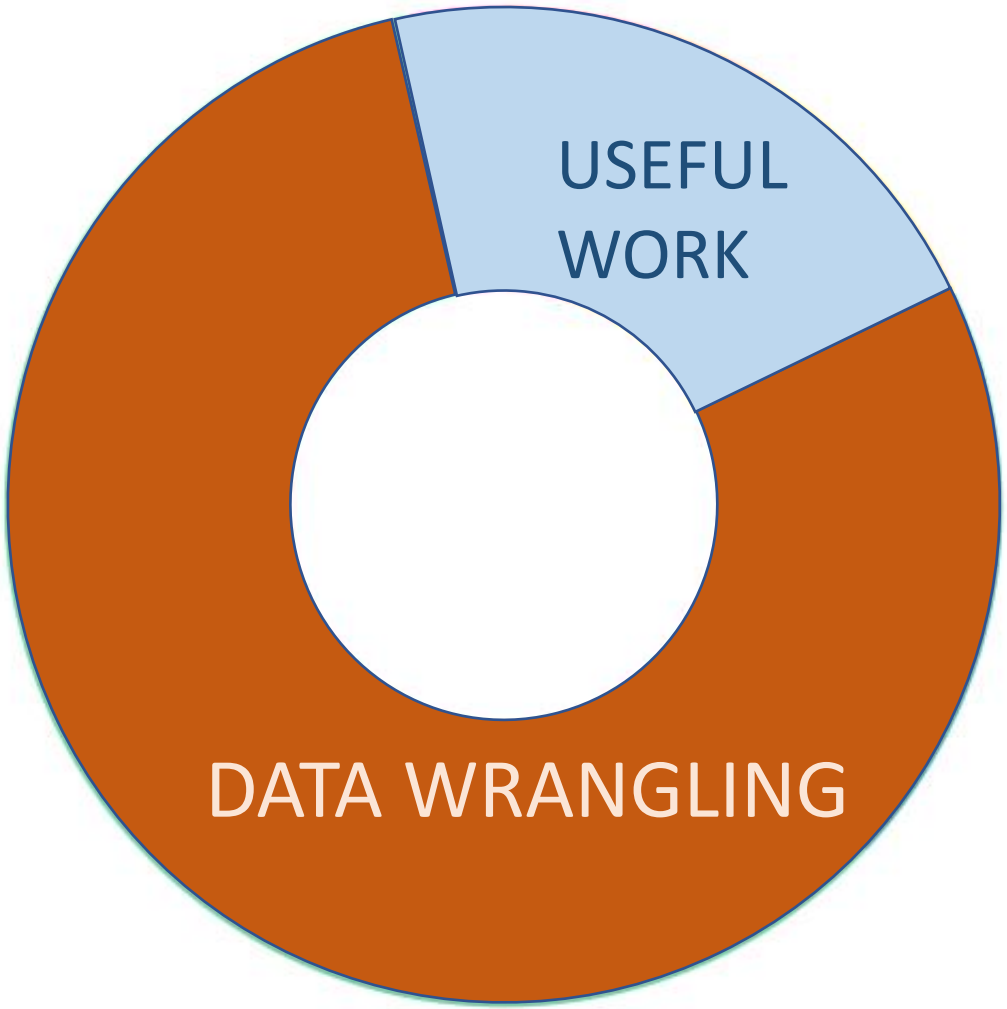
>Analytics

DATA DRIVEN

KNOWLEDGE BASED

DECISION SUPPORT

THE DATA PROBLEM IS UNIVERSAL



What data scientists spend the most time doing

- *Building training sets: 3%*
- *Cleaning and organizing data: 60%*
- *Collecting data sets; 19%*
- *Mining data for patterns: 9%*
- *Refining algorithms: 4%*
- *Other: 5%*

80%

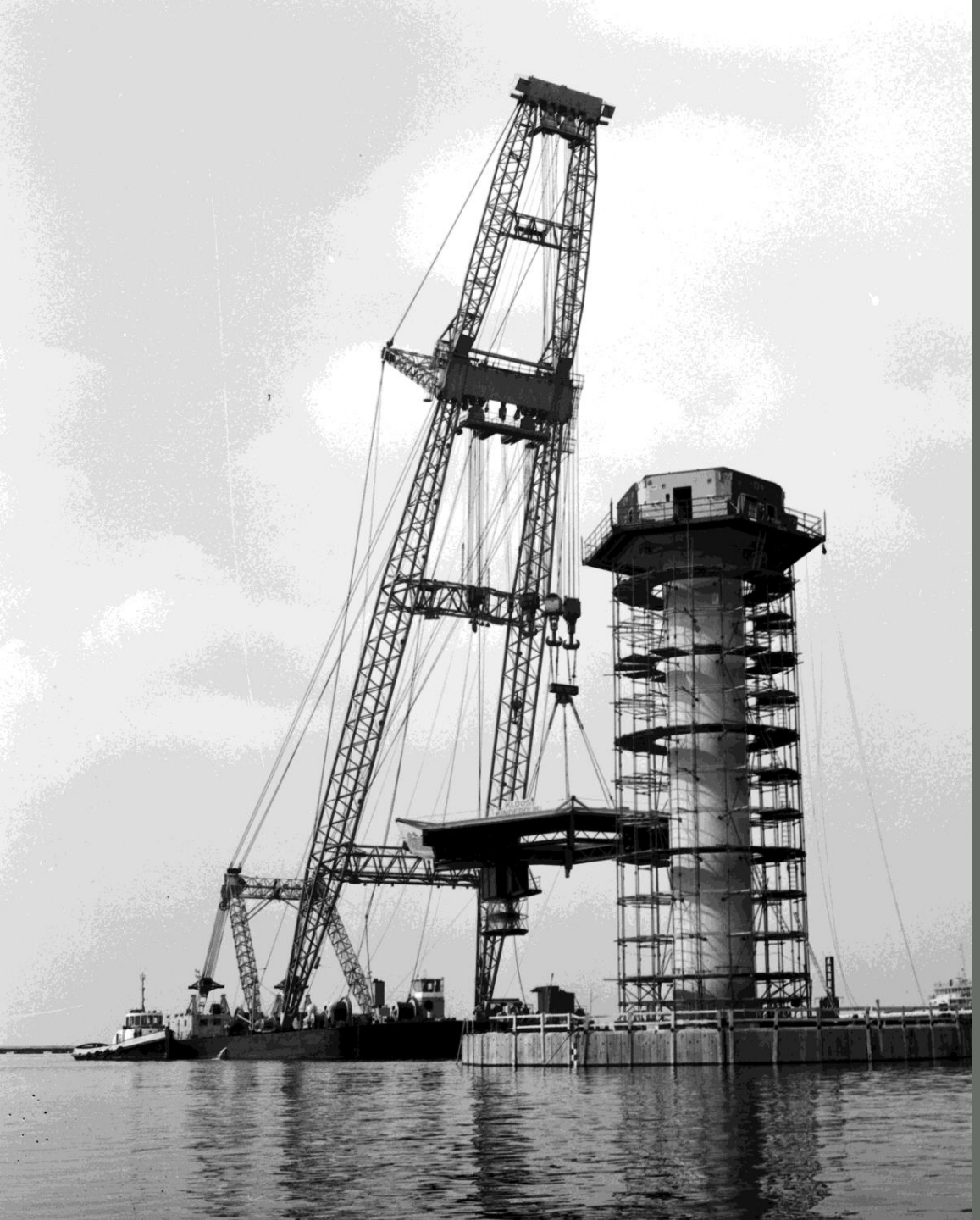
Source: Data Science Report 2016, CrowdFlower, 2016: http://visit.crowdflower.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf

FAIR

Let's aim for:
20% data wrangling
80% useful work

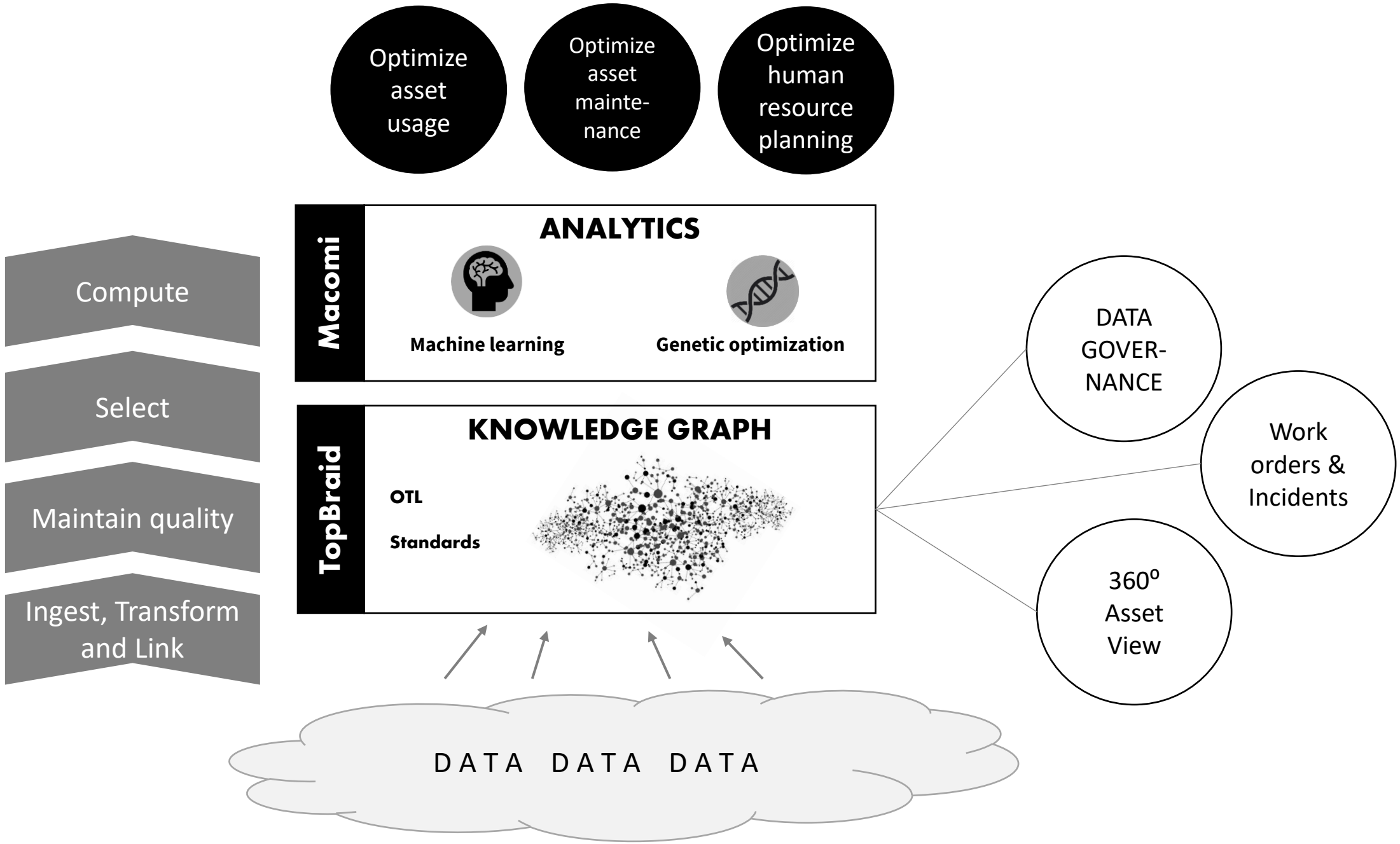
FAIR Principles

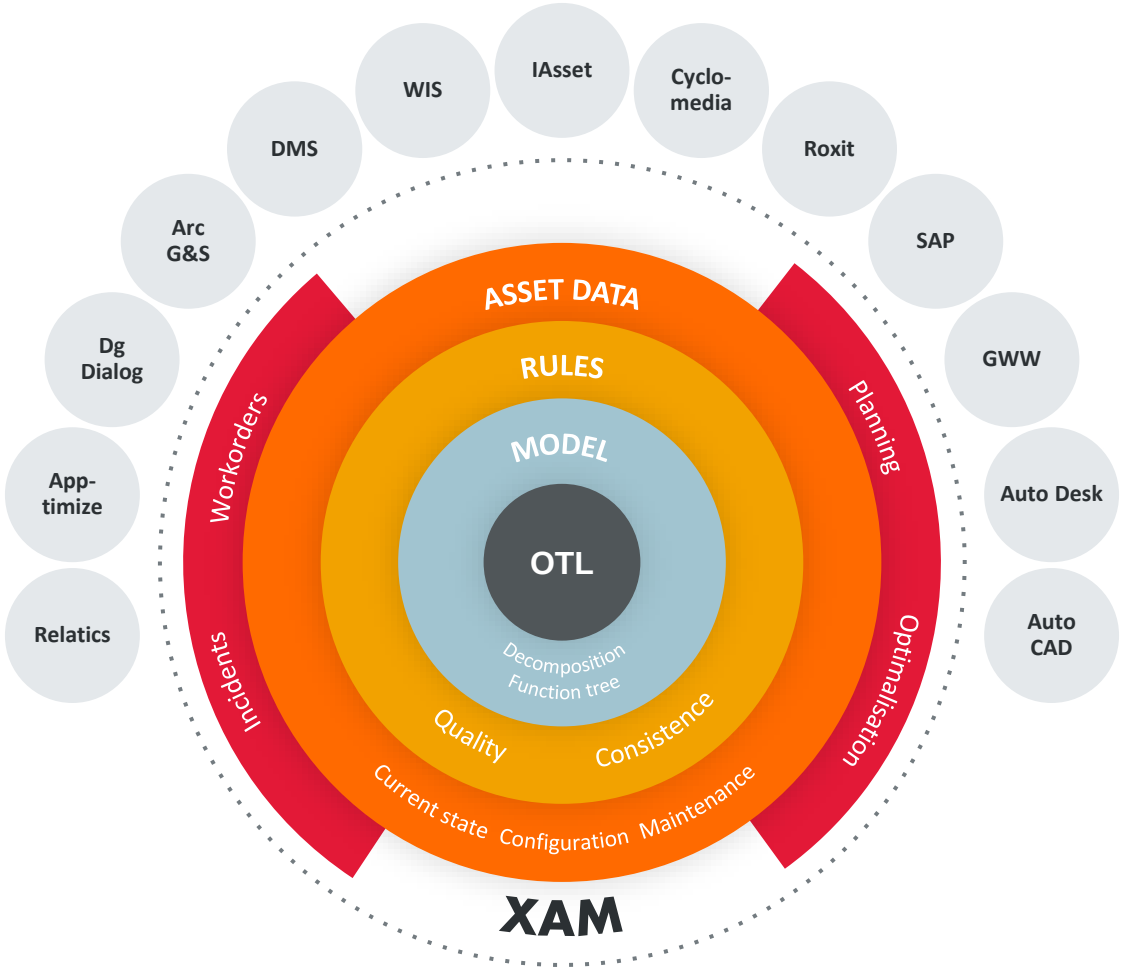
> **F**indable
> **A**ccessible
> **I**nteroperable
> **R**eusable



XAM | ASSET MANAGER

- >Integrated **data** platform
  - ★FAIR principles
  - ★Relevant standards
  - ★Knowledge Graph tech
- >State of the art **analytics**
  - ★Algorithms specific for asset management
  - ★Optimize asset usage, maintenance and resources


Optimize asset usage
Optimize asset mainte-nance
Optimize human resource planning
Compute
Select
Maintain quality
Ingest, Transform and Link
Macomi
ANALYTICS
Machine learning
Genetic optimization
TopBraid
KNOWLEDGE GRAPH
OTL
Standards
DATA GOVER-NANCE
Work orders & Incidents
360º Asset View
DATA DATA DATA


WIS
IAsset
Cyclo-media
DMS
Roxit
Arc G&S
SAP
Dg Dialog
GWW
App-timize
Auto Desk
Relatics
Auto CAD
ASSET DATA
RULES
MODEL
OTL
Decomposition
Function tree
Workorders
Planning
Incidents
Optimalisation
Quality
Consistence
Current state
Configuration
Maintenance
XAM

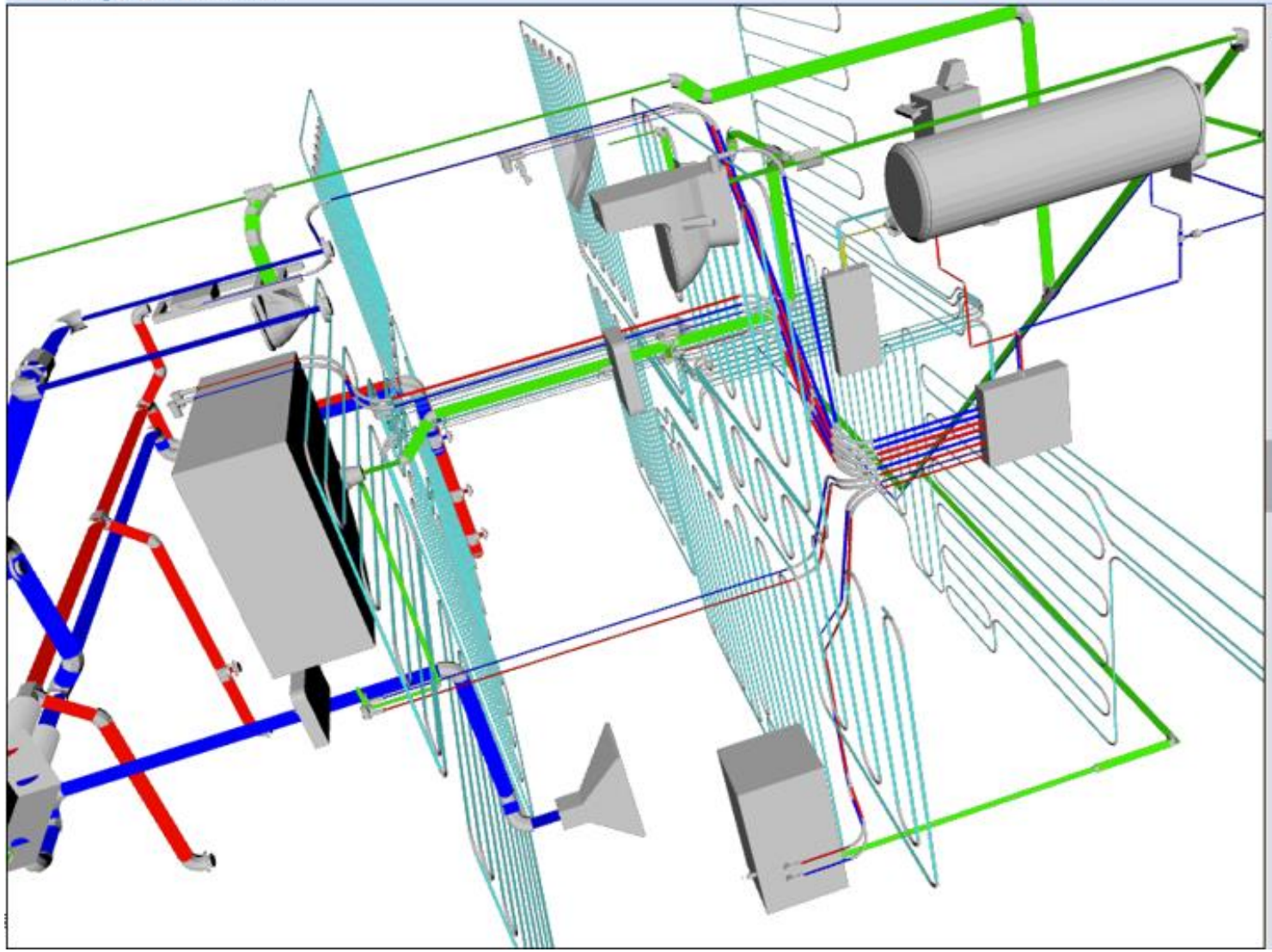
WE ARE HIRING

TopBraid EDG
Enterprise Data Governance

Home > 3D

Global Lookup

User: **David Price** (log out)
Version 5.2.2.v20170116-2101B Server Administration Help



Instance details

urn:x-evn-master:vvs_med_gulvvarme/INSTANCE_711

Annotations

3D DEF: 711

label: EXPRESS instance with name #711 and type IFCENERGYCONVERSIONDEVICE.

Properties

GlobalId: '2his41_kP8Ch2cl2v5FT6y'

ObjectPlacement: EXPRESS instance with name #4522 and type IFCLOCALPLACEMENT.

OwnerHistory: EXPRESS instance with name #75 and type IFCOWNERHISTORY.

Representation: EXPRESS instance with name #4523 and type IFCPRODUCTDEFINITIONSHAPE.

type: IfcEnergyConversionDevice

Incoming References

hasContents (inverse): c9c199-6ea6-47da-940a-4abece14ebc9
cbd7b04-a8d7-4244-9863-829e85066526
f57b961-cc81-4cd7-932b-ae566ced48a8
f7829a2-0250-4daf-91fe-220378a11010
fb8b07-b8a8-46bb-9e95-5b631416b67c

PERSPECTIVES

Asset Owner Perspective

Contractor Perspective

Schiphol Data Dictionary
OTL

Customer's dataplatform

Waternet
OTL

Customer's dataplatform

Customer X
OTL

Customer's dataplatform

Asset Info Viewer

APPS

Datahub

THE DATA HUB

Some more details

Data Quality

360° Asset View

DATA ENGINEER

END USER

EDG

TopBraid Enterprise Data Governance™

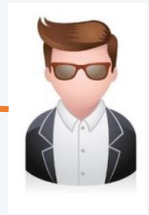
EXPLORER | API

TopBraid Enterprise Data Governance™

{ REST:API }

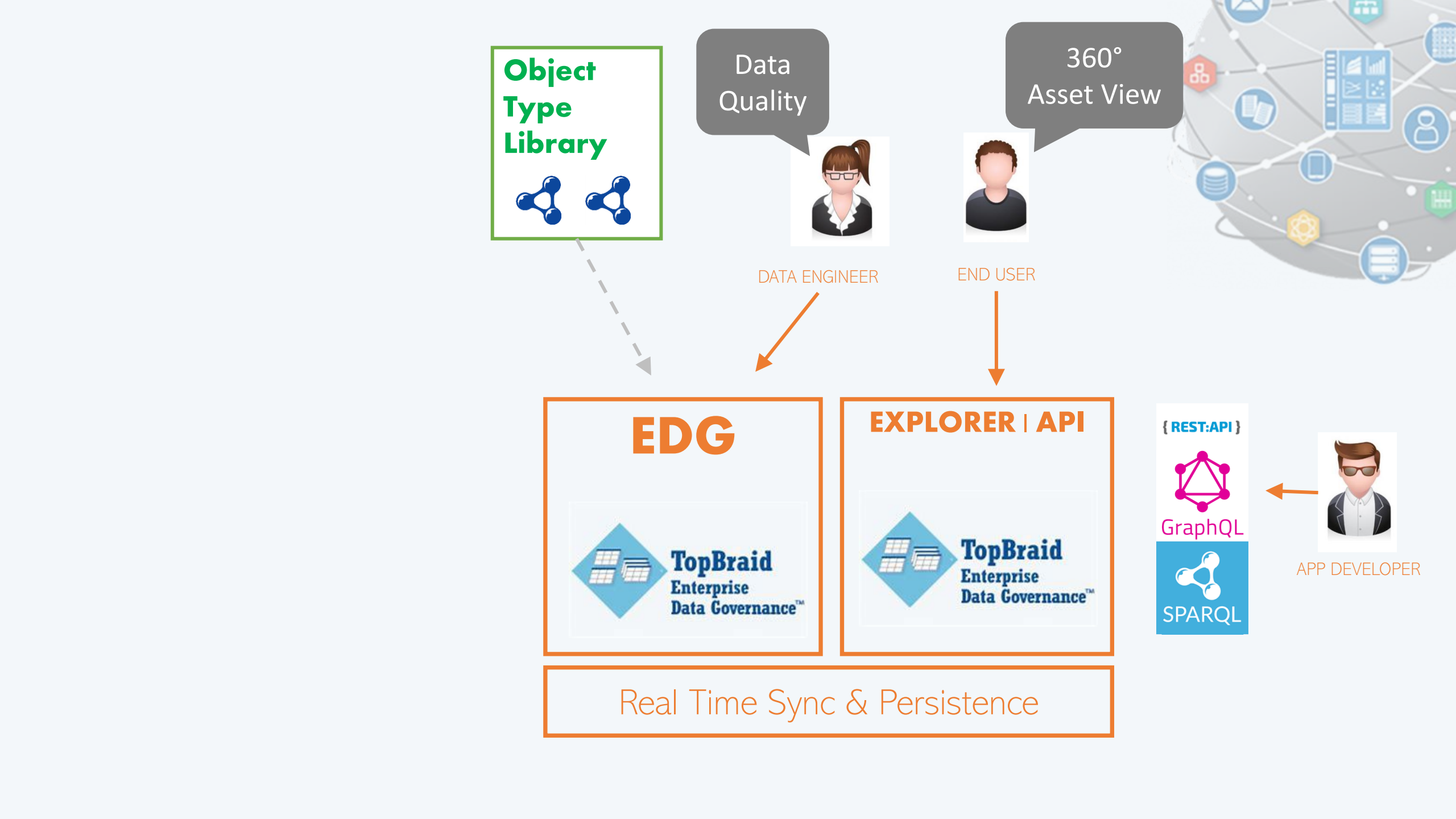
GraphQL

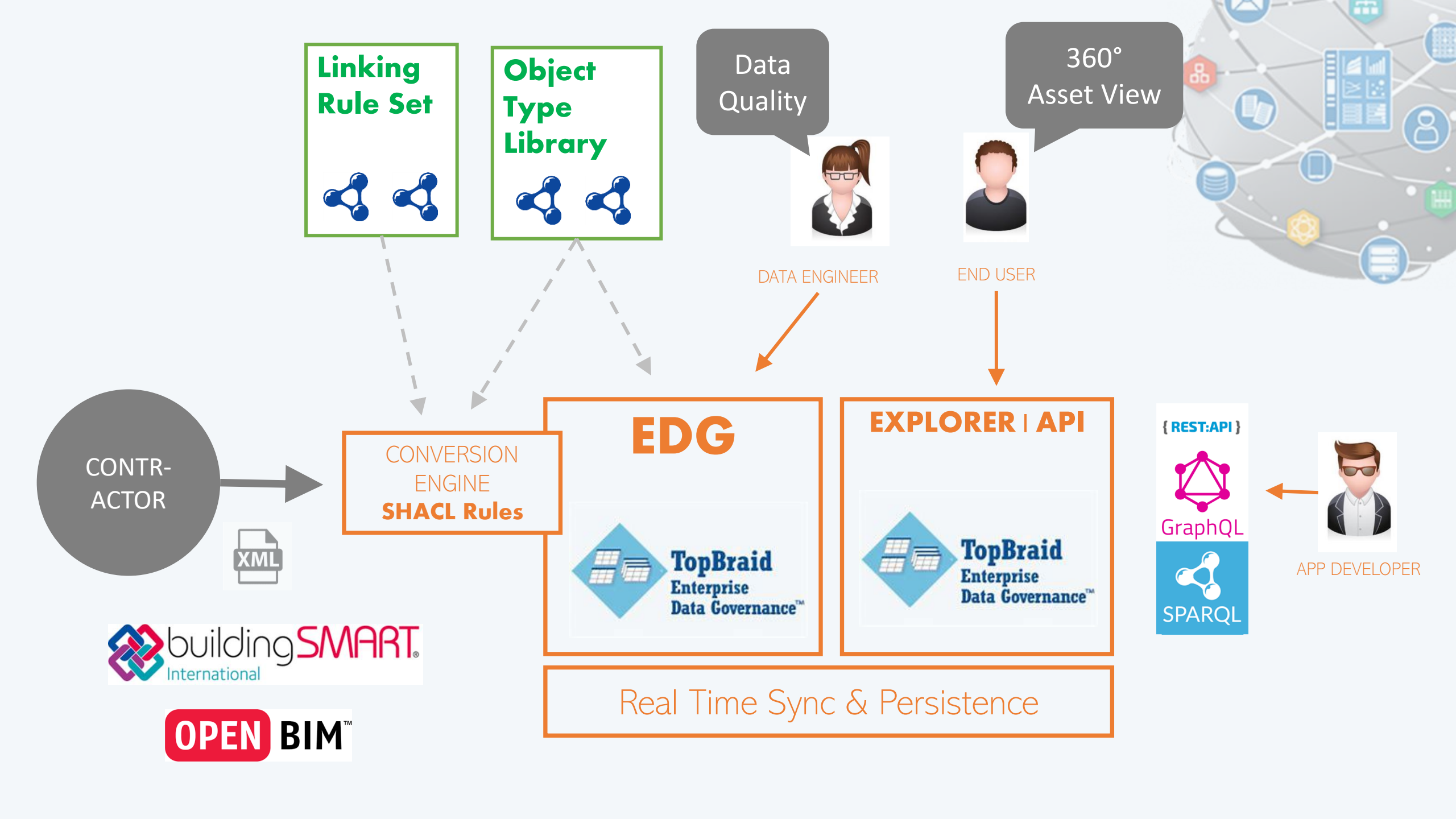
SPARQL



APP DEVELOPER

Real Time Sync & Persistence


Object Type Library
Data Quality
360° Asset View
DATA ENGINEER
END USER
EDG
TopBraid Enterprise Data Governance™
EXPLORER | API
TopBraid Enterprise Data Governance™
{ REST:API }
GraphQL
SPARQL
APP DEVELOPER
Real Time Sync & Persistence


Linking Rule Set
Object Type Library
Data Quality
360° Asset View
DATA ENGINEER
END USER
CONTR-ACTOR
XML
CONVERSION ENGINE
SHACL Rules
EDG
TopBraid Enterprise Data Governance™
EXPLORER | API
TopBraid Enterprise Data Governance™
{ REST:API }
GraphQL
SPARQL
APP DEVELOPER
buildingSMART International
OPEN BIM™
Real Time Sync & Persistence

STANDARDS IN THE MAKING

NETHERLANDS

COINS [OWL + proprietary interpretation]
NEN NTA 8035 [OWL, SHACL]
GWSW [OTL for urban water management]
BGT [LOD and GIS combined (Kadaster)]
IMBOR [Management of Public Space]

OTLs published by:
RWS, ProRail, Waternet ... (many will follow, soon)

INTERNATIONALLY

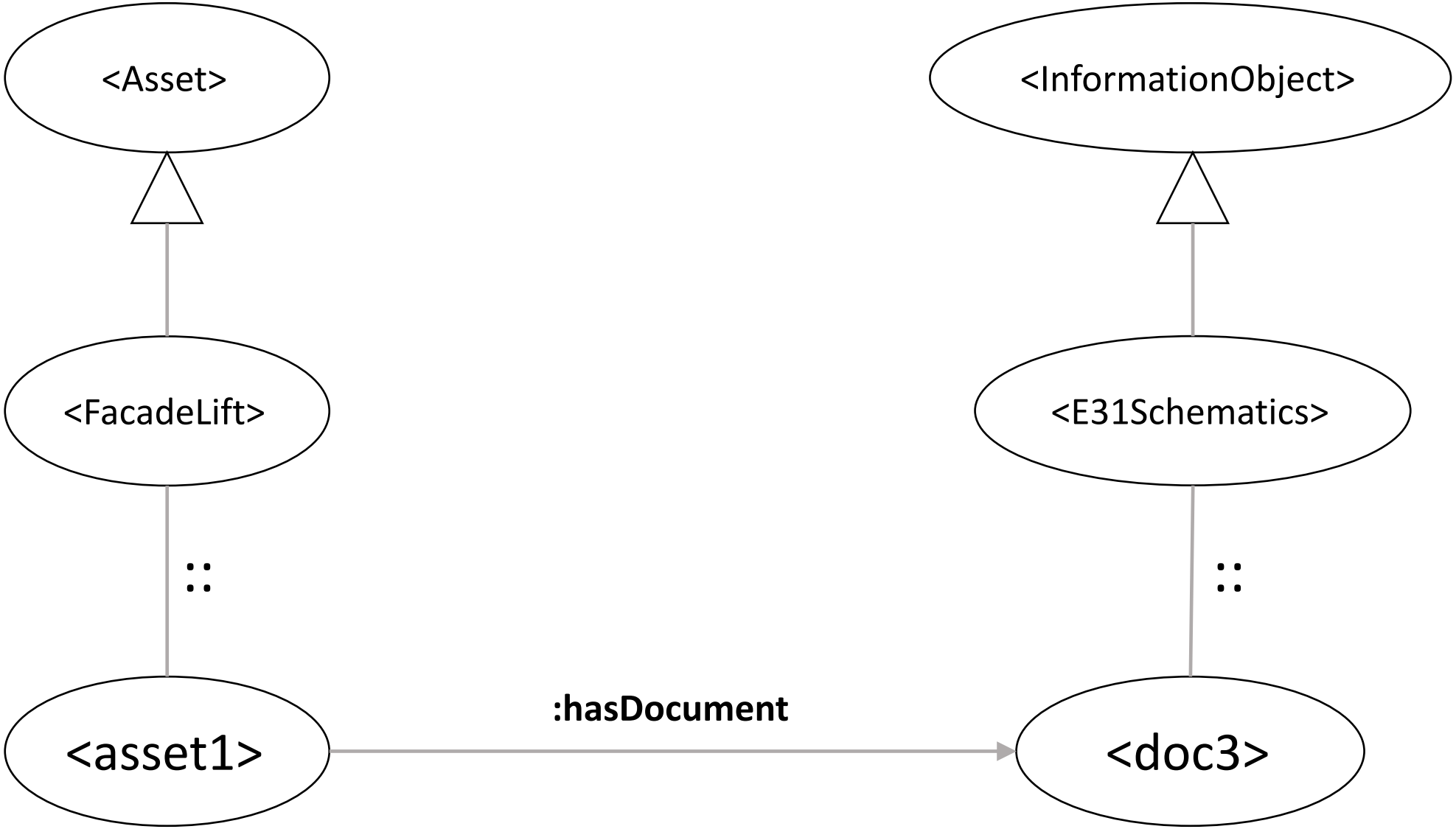
SAREF4BLDG [SAREF]
Linked Building Data (BOT, ...) [W3C]
buildingSMART [IFC-OWL > ¿native RDF?]
CEN TC442
ISO 21597 ICDD

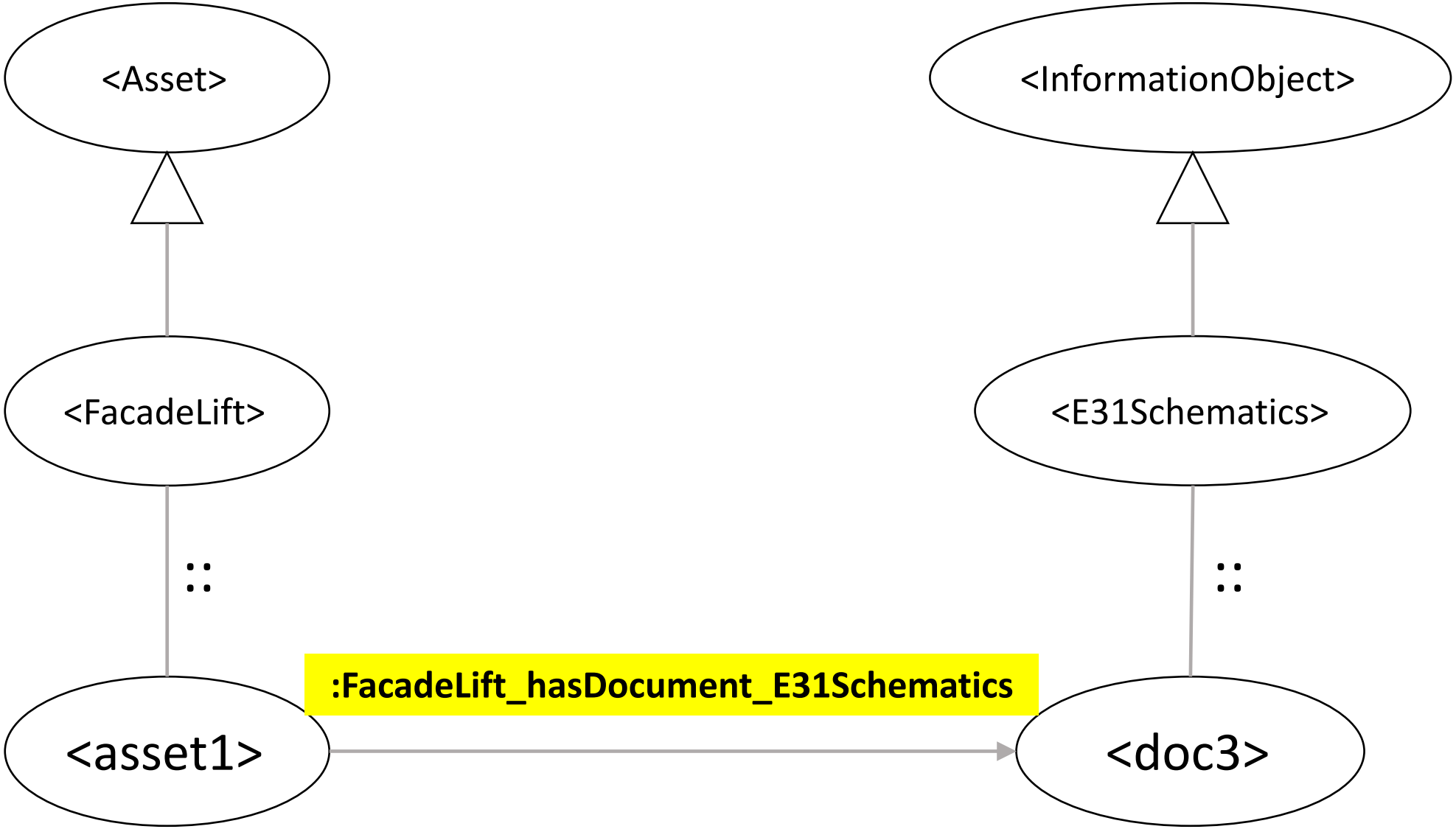
CHALLENGES

Ontology anti-patterns

People and parties creating OTLs have very different backgrounds

- Unnecessary inverse properties (see PROV-O for interesting discussion)
- Overusing OWL
- Underusing SHACL
- L1 vs L2 vs L3-modelling styles
- Conceptual modelling errors (see the ontoUML-community, Guizzardi 2006)
- … and the list goes on


<Asset>
<InformationObject>
<FacadeLift>
<E31Schematics>
::
::
<asset1>
:hasDocument
<doc3>


<Asset>
<InformationObject>
<FacadeLift>
<E31Schematics>
::
::
:FacadeLift_hasDocument_E31Schematics
<asset1>
<doc3>

QUERY:

"The average number of documents per asset instance in the database"

OLD SITUATION:

70,000 instances of asset x 1

Solution space: 70,000
QUERY TIME: 0.1 seconds

NEW SITUATION:

70,000 instances of asset x
4498 solutions for "subproperty of hasDocument"

Solution space: 314,860,000
QUERY TIME: 7.5 minutes

THE OTL / ONTOLOGY IS NOT JUST A DATA EXCHANGE FORMAT

IT IS THE "SOURCE CODE" THAT CAUSES THE DATA TO BEHAVE AS THEY DO

HOW DO WE MAKE SURE OTL'S ARE FAIR?

Your participation is needed!

THANK YOU!

Jan Voskuil

jan.voskuil@taxonic.com

TAXONIC