



Mass Balance Models, Triples, and SPARQLs

November 06, 2018

Kostadin Alargov, Matthew Kramer, Siping Wang (TetraScience)
Benjamin Woolford-Lim, James Roberts, Joy Ginocchio (GSK)

Gartner

Cool
Vendor
2017

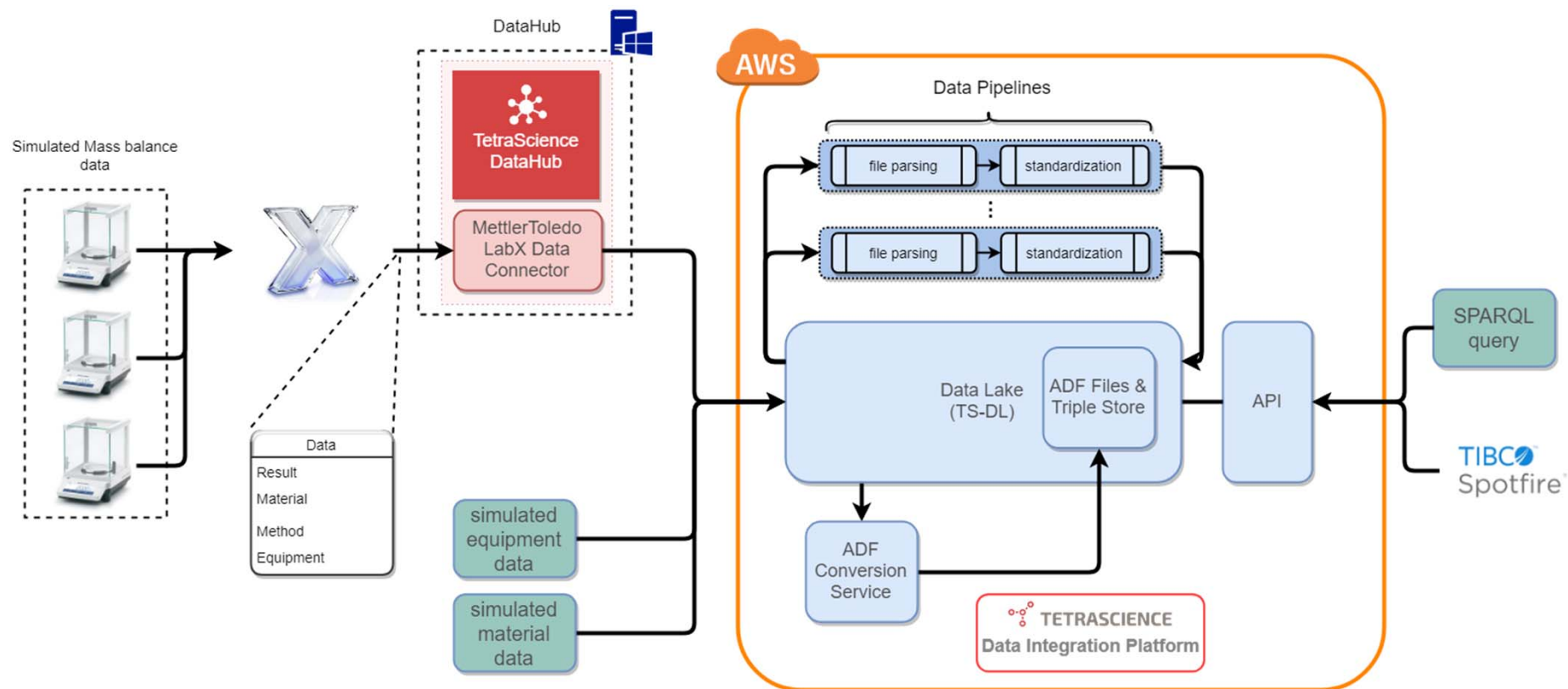


Motivation

- A new way to explore and interact with ADF for the Community
 - Triple Store graph database
- Support ongoing member company efforts
 - GSK Integration Project to support First-Time in Human analytical processes
- Expand current provisional data models and link data from different domains
 - Mass Balance
 - Equipment
 - Material
- Stress test the semantic concepts used in data models
 - SPARQL queries

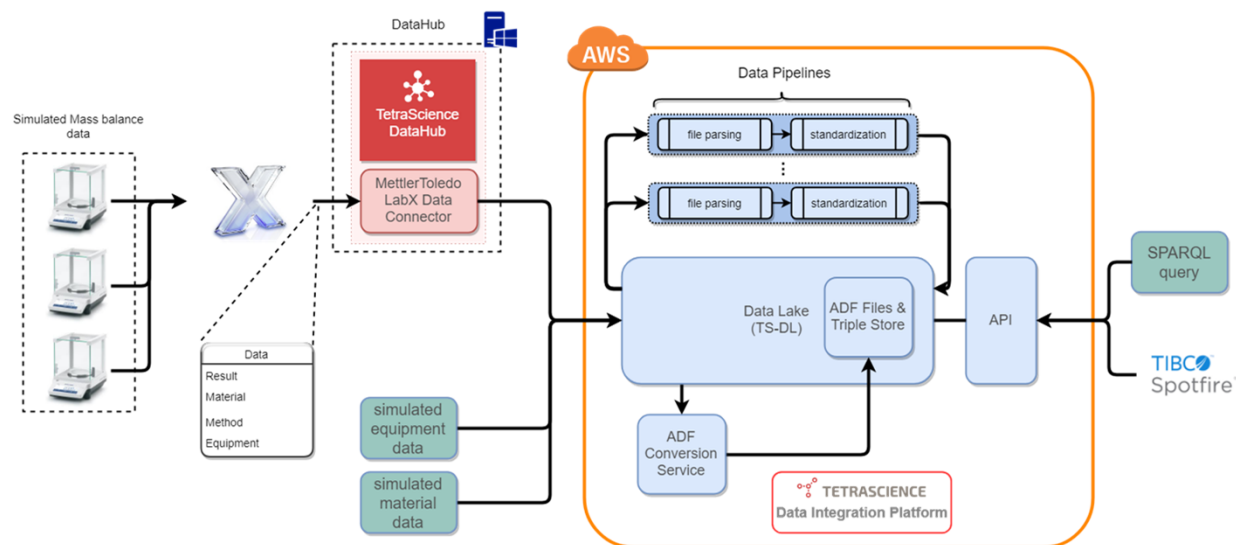
Architecture

3



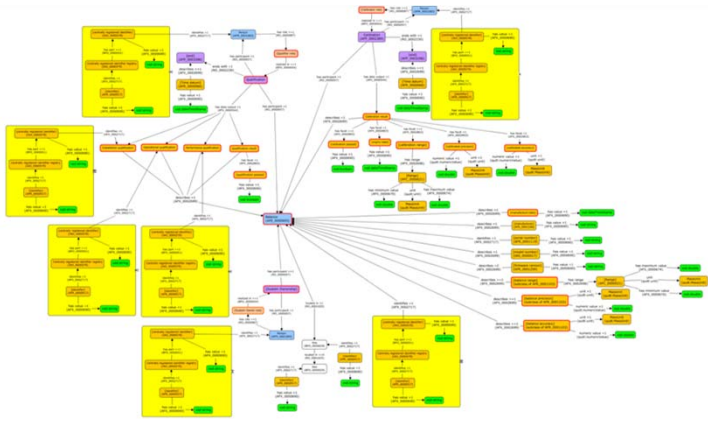
Architecture

- Collect data from sources (e.g. LabX)
- Transfer to cloud infrastructure
- Convert to ADF
- Create Graph Database (Triple store)

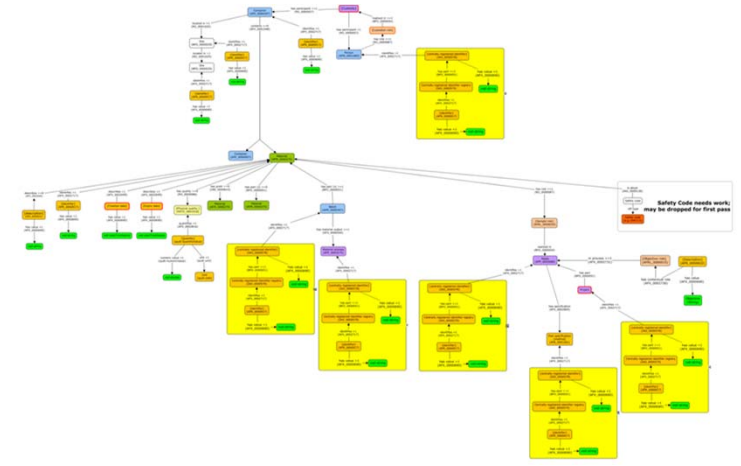


Provisional Models

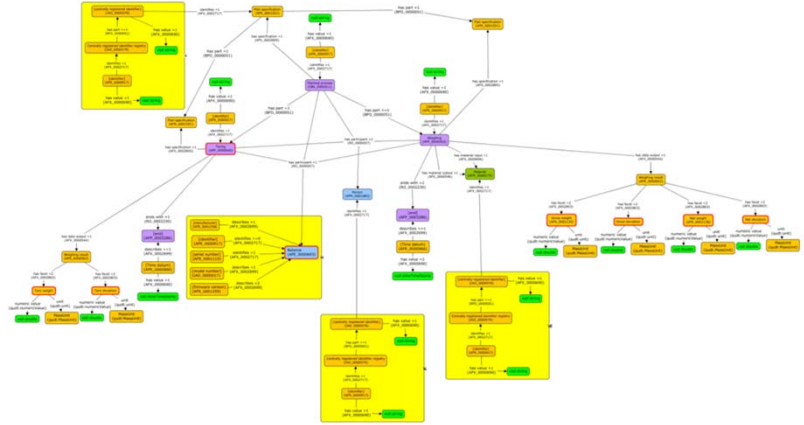
Device Model



Material Model

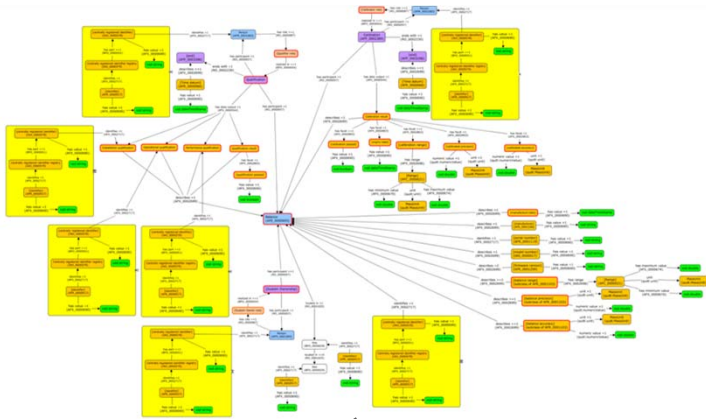


Weighing model

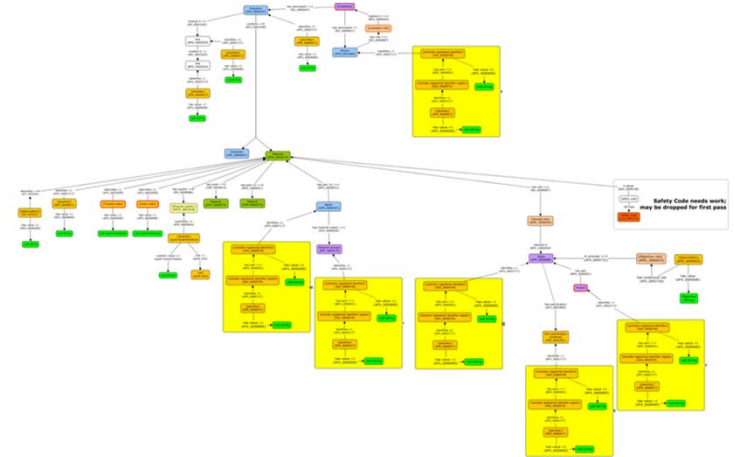


Provisional Models

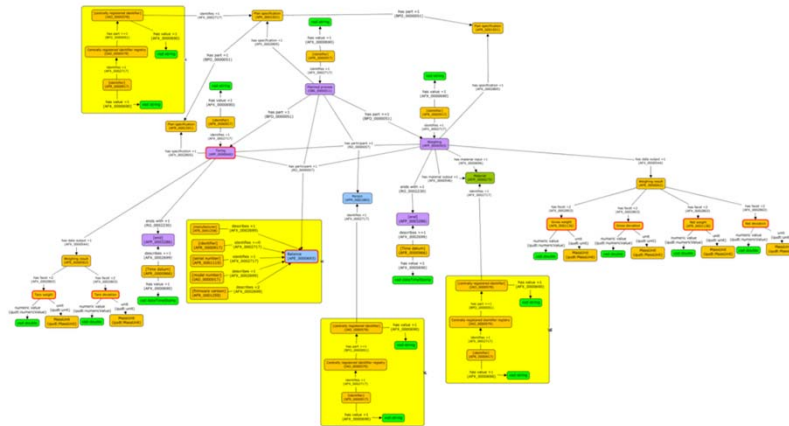
Device Model



Material Model



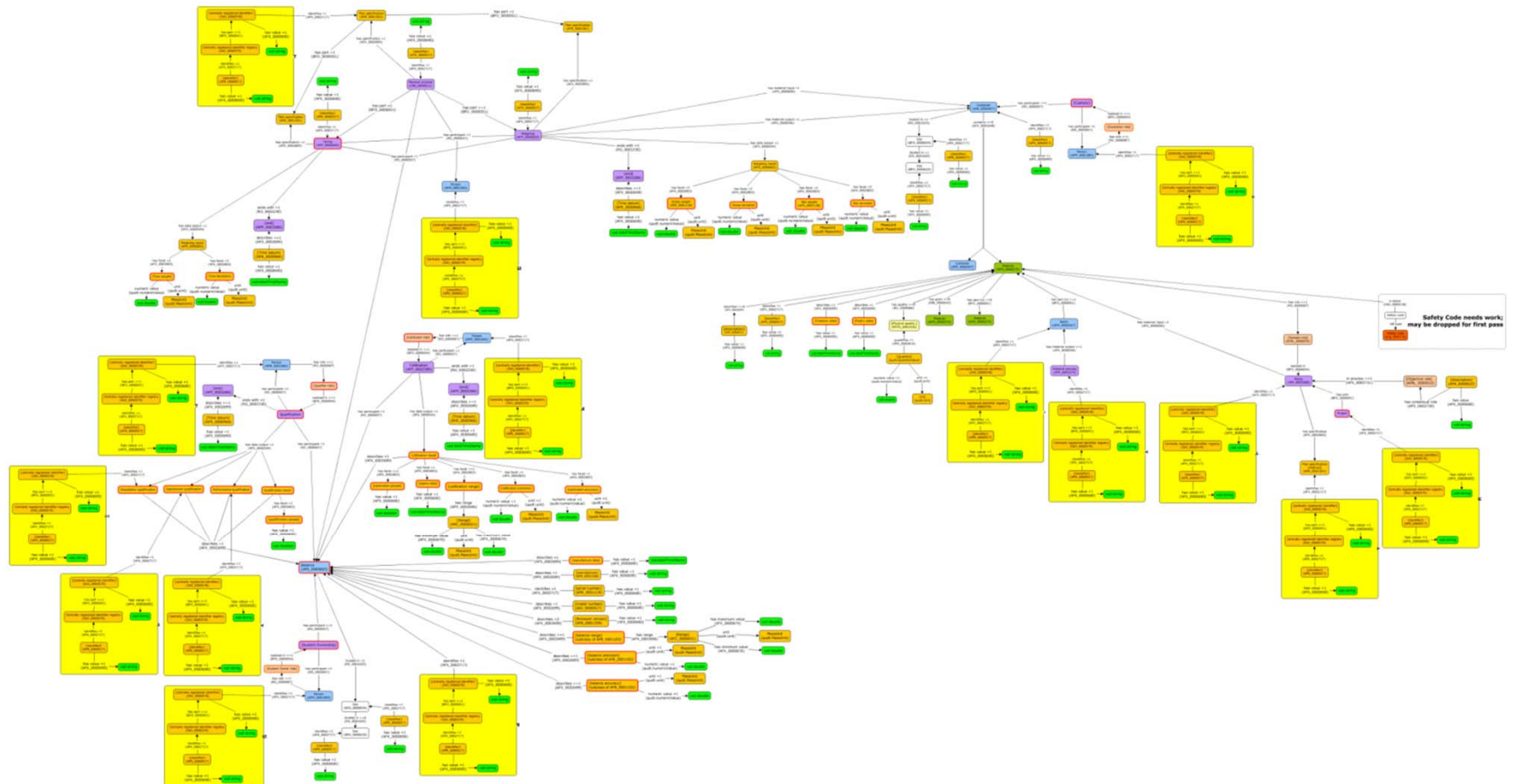
Weighing model





Provisional Merged Models

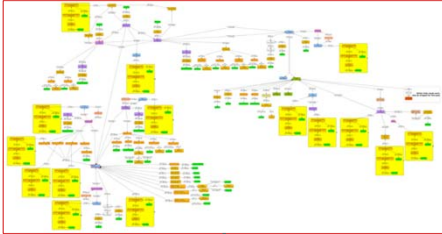
7



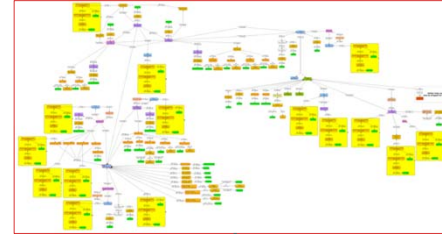


Merged Models

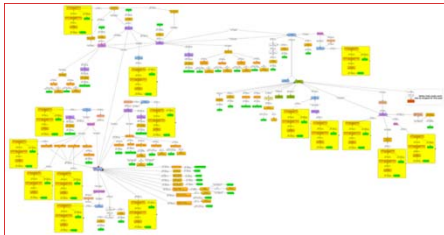
Instance data 1 (ADF)



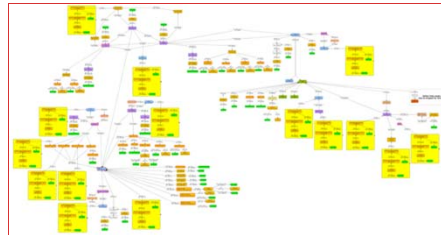
Instance data 1 (ADF)



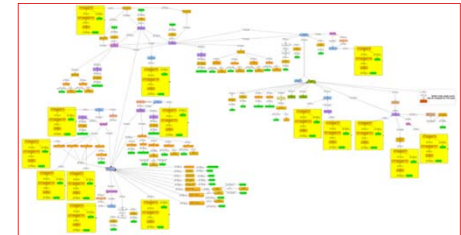
Instance data 1 (ADF)



Instance data 1 (ADF)



Instance data 1 (ADF)

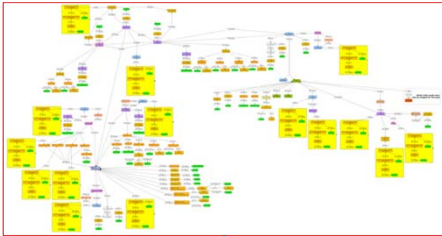


Triple Store Graph
Database

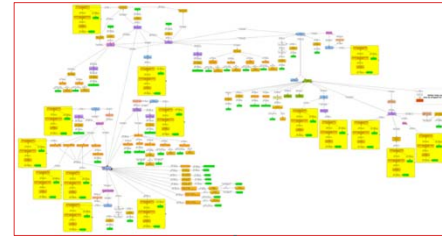


Merged Models

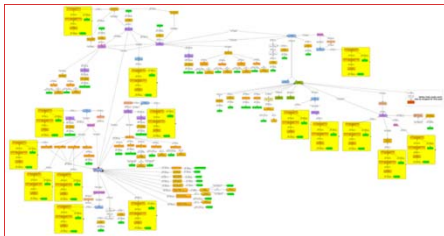
Instance data 1 (ADF)



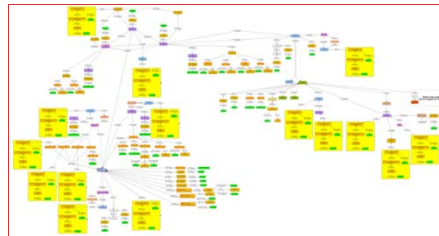
Instance data 1 (ADF)



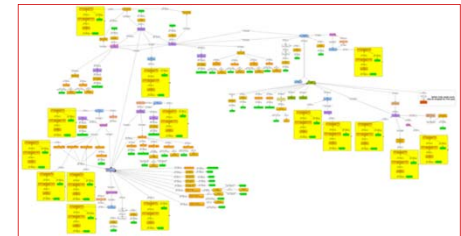
Instance data 1 (ADF)



Instance data 1 (ADF)



Instance data 1 (ADF)



Triple Store Graph
Database

SPARQL Queries



User Stories...

1. **As a metrology specialist, I need to see which balances require calibration in a given time frame so I can schedule time to calibrate them and ensure they remain compliant.**
2. **As a lead investigator carrying out a quality investigation, I need to know all of the weighing events that were performed on a particular balance in a particular date range so I can complete an impact assessment.**
3. **As a downstream software agent, I need to be able to query for the results for a given container ID (i.e. barcode) to be able to complete my downstream actions (e.g. visualizations, calculations).**



...into SPARQL queries...

#weighing

```
?weighing obo:RO_0000057 ?balance .

?weighing rdf:type/rdfs:subClassOf* af-p:AFP_0000503 .

?weighing obo:RO_0002230 ?weighingEnd .

?weighingEndTimeDatum af-x:AFX_0002699 ?weighingEnd .

?weighingEndTimeDatum af-x:AFX_0000690 ?weighingEndTimeStamp .
```

#material ID

```
?weighing af-x:AFX_0000696 ?container .

?container rdf:type/rdfs:subClassOf* af-e:AFE_0000407 .

?container af-x:AFX_0001048 ?material .

?material rdf:type/rdfs:subClassOf* af-m:AFM_0000275 .

?materialIdentifier af-x:AFX_0002717 ?material .

?materialIdentifier rdf:type/rdfs:subClassOf* obo:IAO_0000578 .

?materialIdentifier af-x:AFX_0000690 ?materialId .
```

#material registry ID

```
?materialRegistry obo:BFO_0000051 ?materialIdentifier .

?materialRegistry rdf:type/rdfs:subClassOf* obo:IAO_0000579 .

?materialRegistryIdentifier af-x:AFX_0002717 ?materialRegistry .

?materialRegistryIdentifier rdf:type/rdfs:subClassOf* af-r:AFR_0000917 .

?materialRegistryIdentifier af-x:AFX_0000690 ?materialRegistryId .
```



...at your fingertips



In Summary

- Working, provisional semantic data models can be developed quickly: the merged data model described here was developed in about 1 week
- Simulated instance data is a fast and cost-effective way to explore and refine a data model
- The Allotrope Community will be granted free access to the triplestore and Spotfire visualizations so we can all test the model and queries on scale (timing and duration to be announced in 2019)



Acknowledgement

TetraScience Team

- Milos Grbic
- Vincent Chan
- Nenad Vukicevic



Thank You!!!



The Gartner Cool Vendor Logo is a trademark and service mark of Gartner, Inc., and/or its affiliates, and is used herein with permission. All rights reserved. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.