Building and exploring Knowledge Graphs with eccenca Corporate Memory

Dr. Natanael Arndt
Senior Linked Data Expert @ eccenca GmbH

Leipzig

ASCM - APICS

Business Digital Twin

BOSC

EDM Daimler

BERLIN

Council Innovate with Data

>100% Growth

NOKI

Volkswagen

60 FTE

Knowledge Graph

VOITH

DBpedia

DaaS/Cloud

Gartner

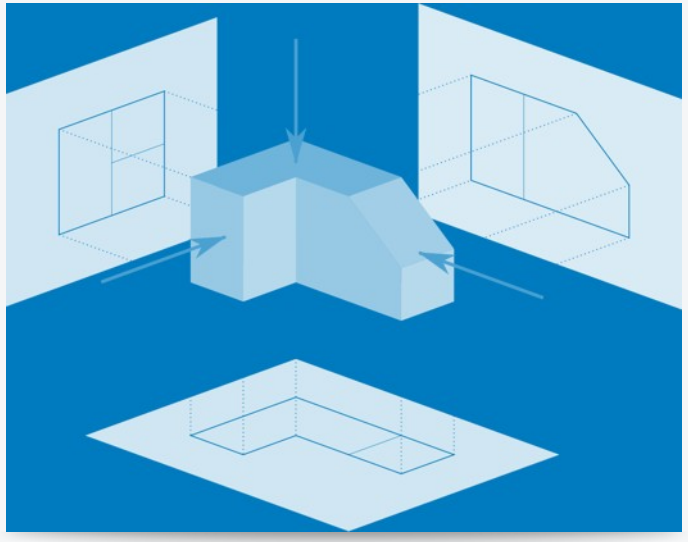
Cool Vendor 2018

Building and exploring Knowledge Graphs with

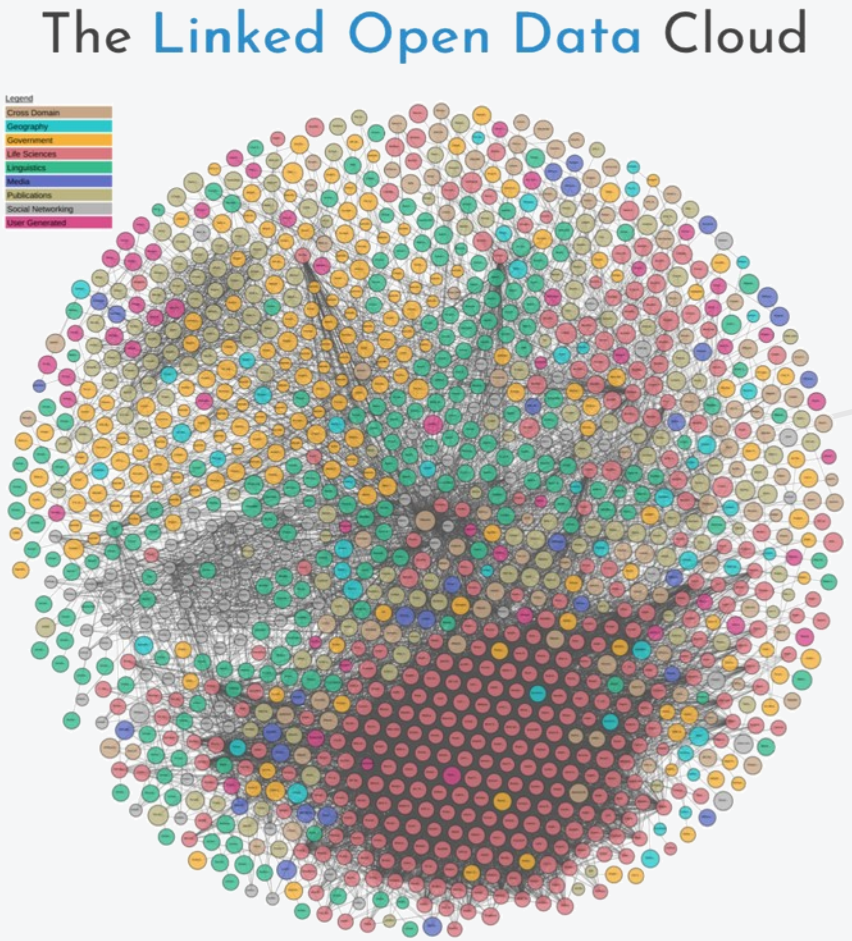
eccenca Corporate Memory

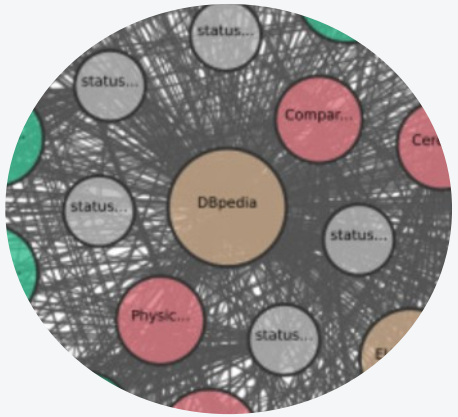
Perspectives on data turn into silos

Parts of data are being curated, duplicated, annotated and simply changed over time, making reconciliation and interpretation a challenge



Data on the web & WikiData & DBpedia
Global and Unified Access to Knowledge



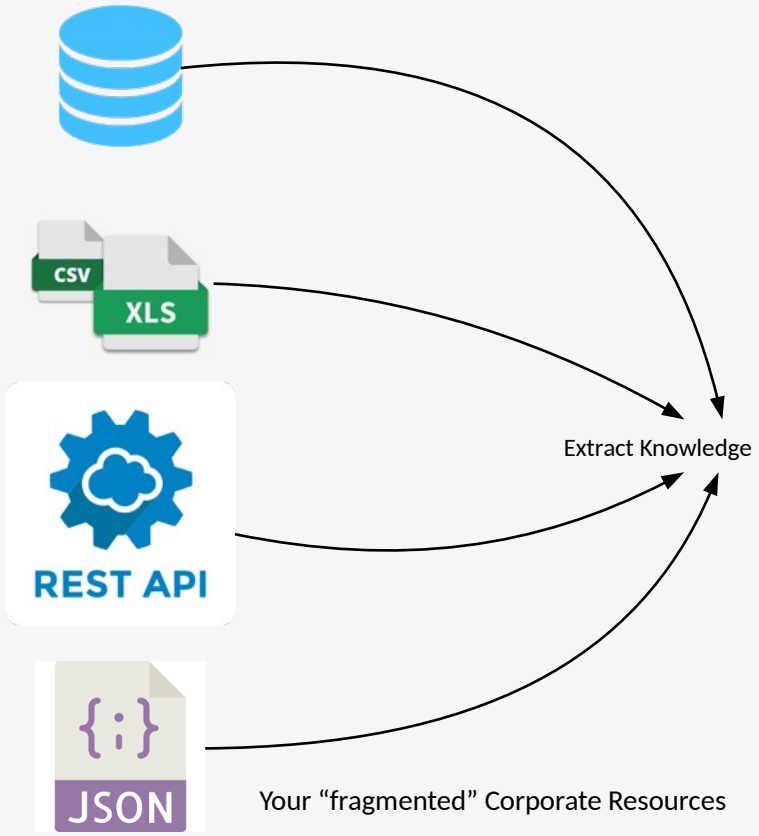


DBpedia extracts structured content from Wikipedia.

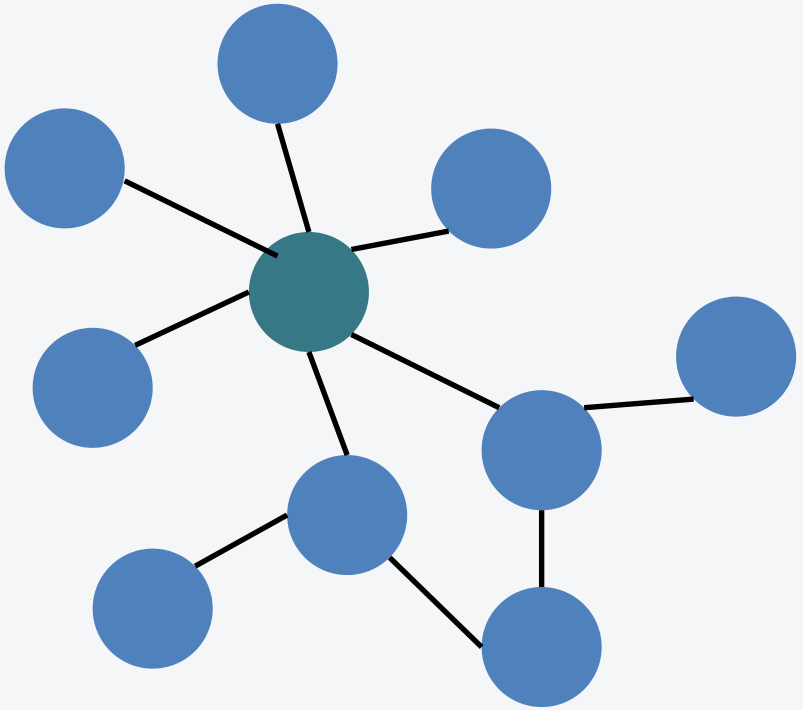
This structured information is made available on the World Wide Web.

Leipzig University - Prof. Dr. Sören Auer, co-founder of Eccenca

From the web to linked enterprise data
Global & unified access to corporate knowledge

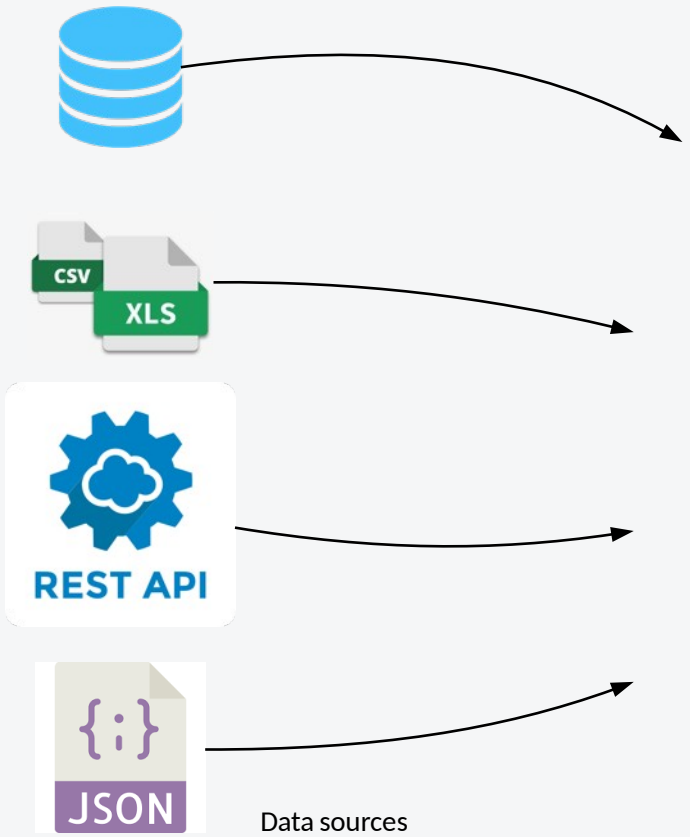


Your “fragmented” Corporate Resources

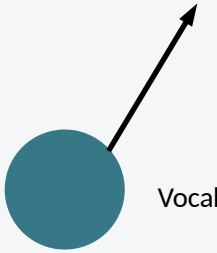


What is your Corporate “DBpedia” ?

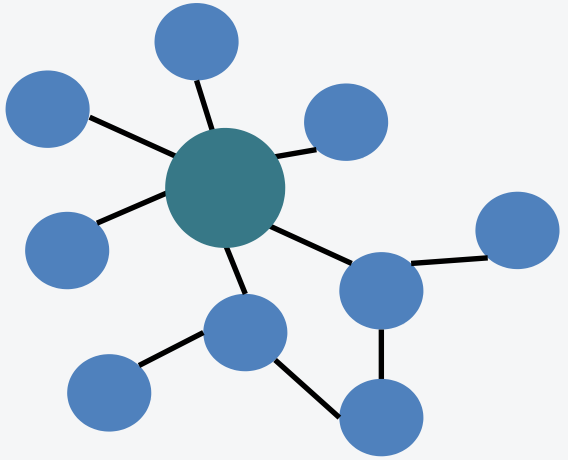
eccenca Corporate Memory



Data sources

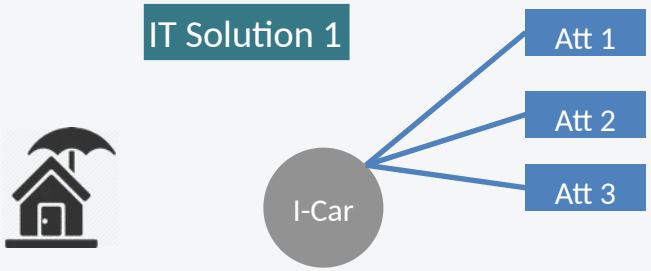


Vocabularies



eccenca Knowledge Graphs

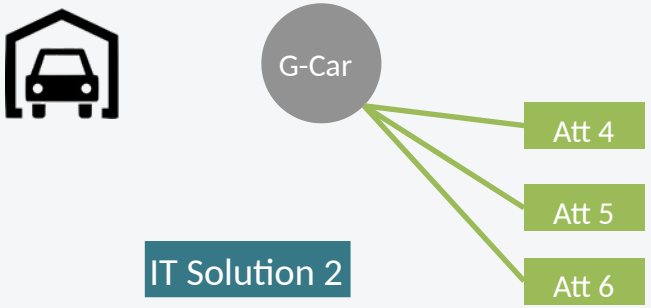
Global identification, multiple perspectives

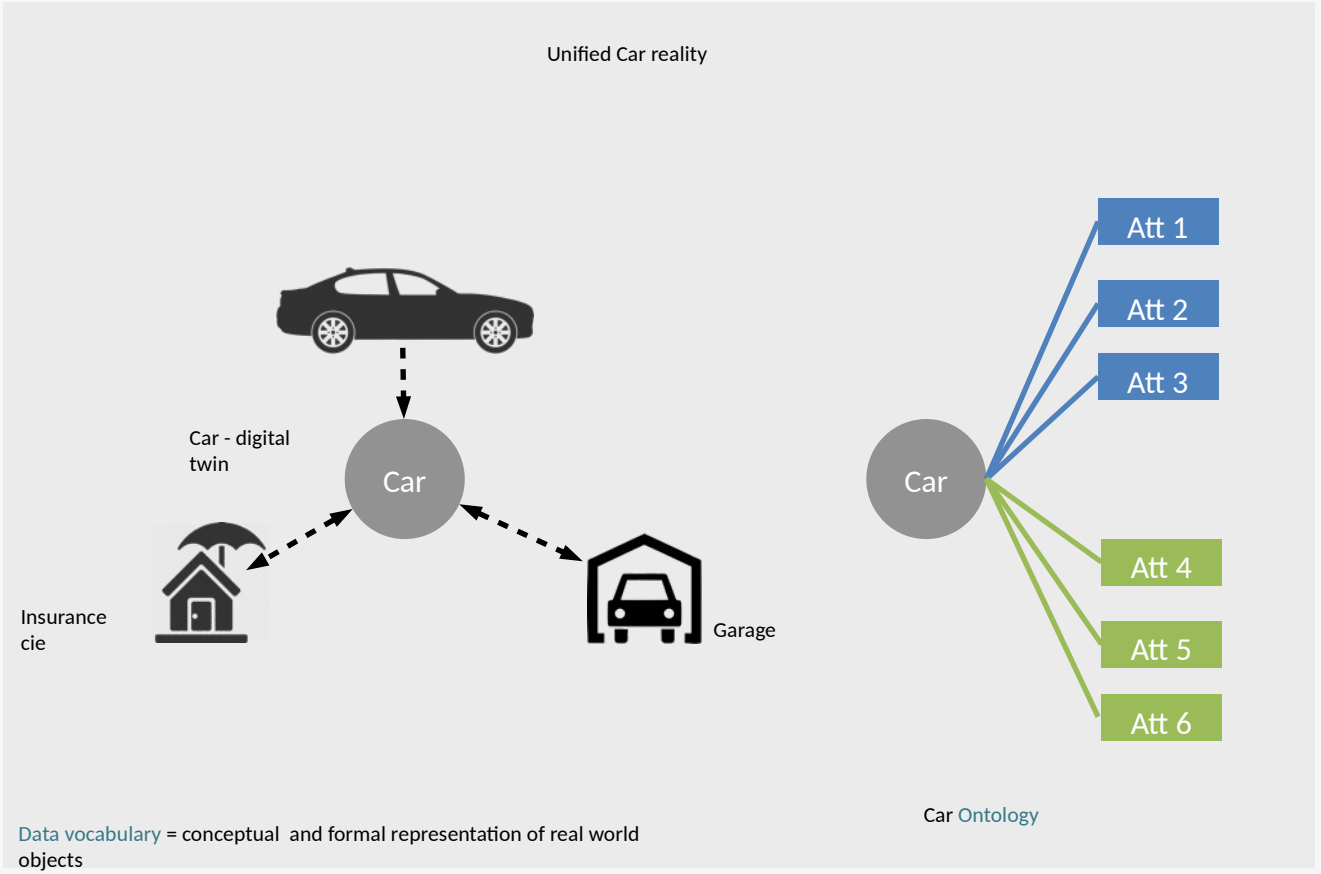


Insurance Car reality

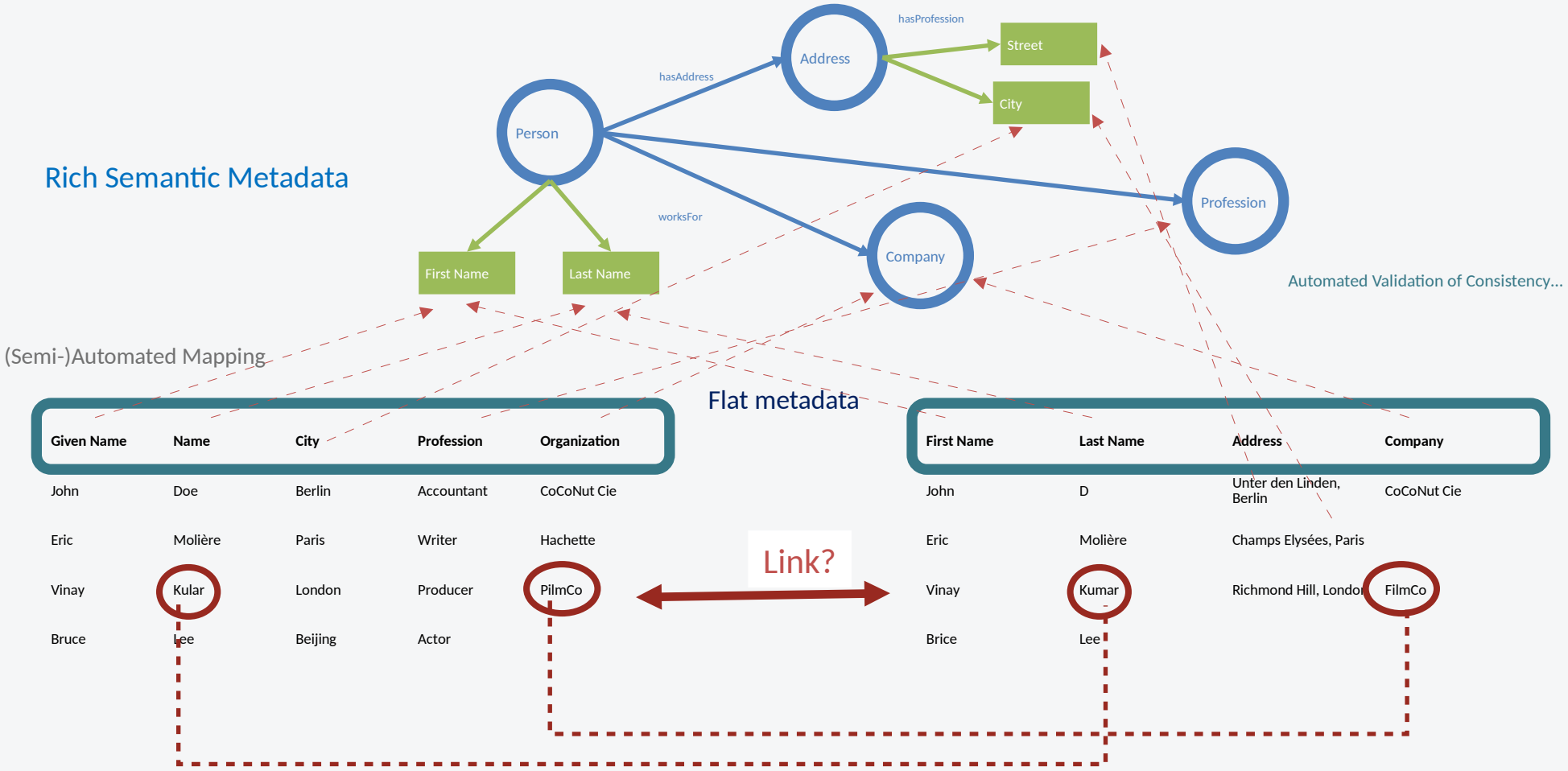
Application-level Integration

Garage Car reality

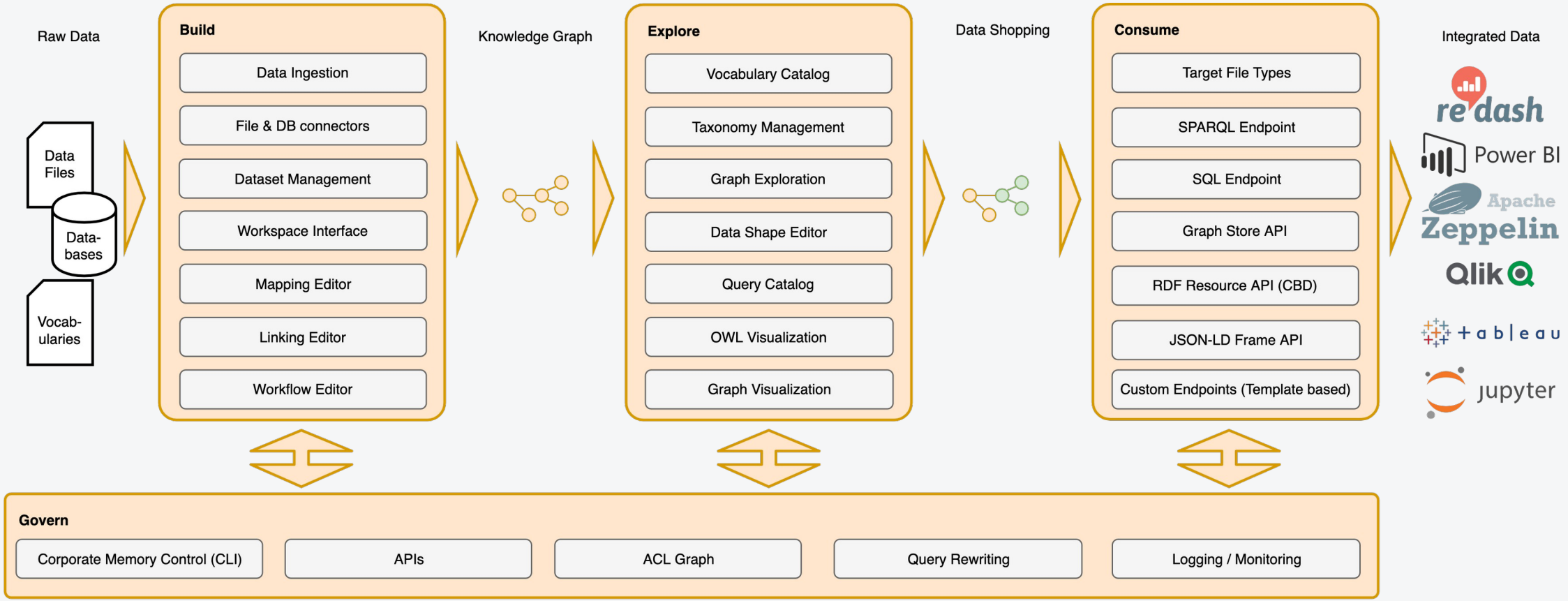




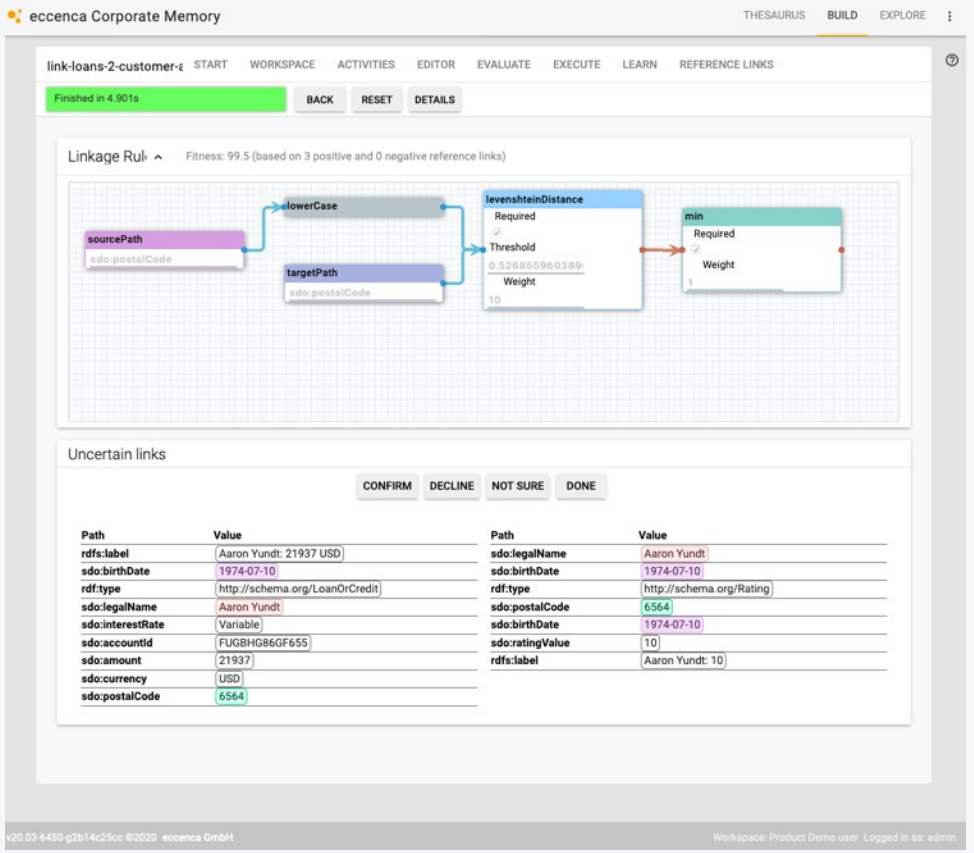
Semantic metadata vs flat metadata

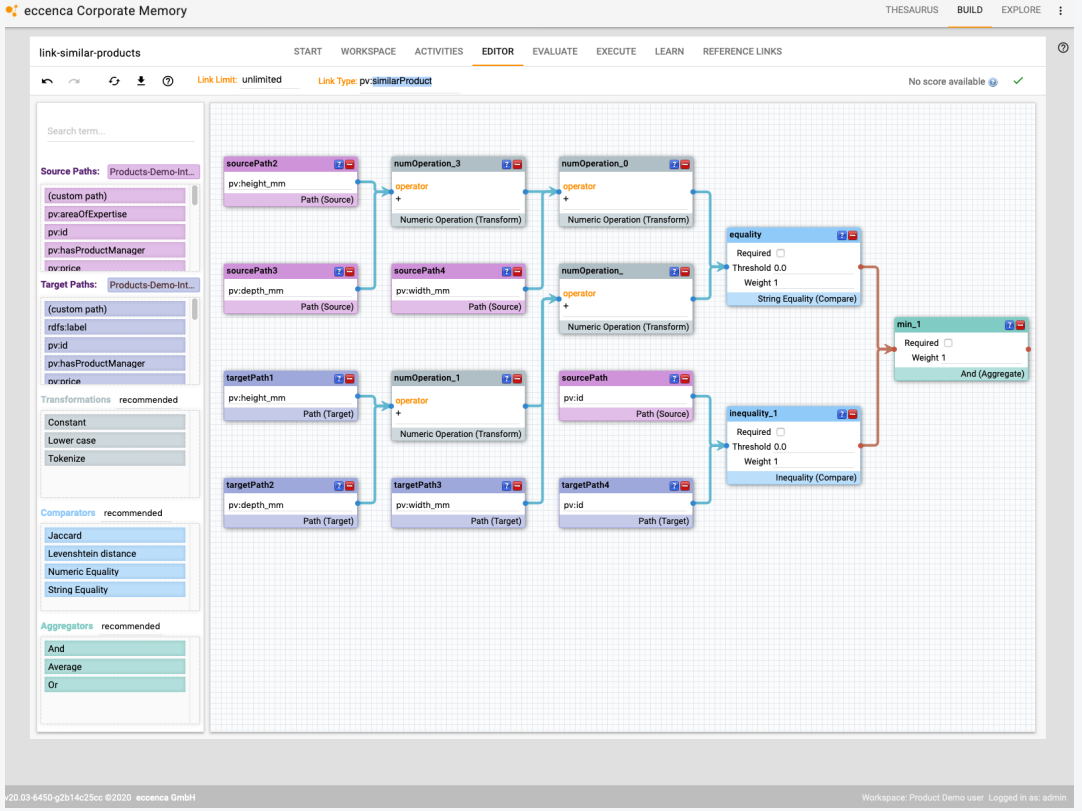


User Journey and Functional Areas



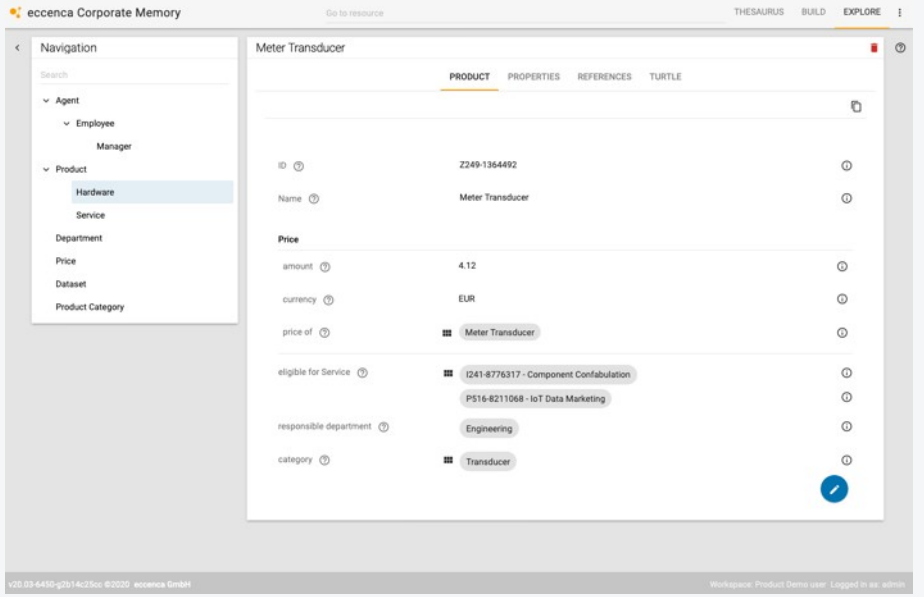
Linking as general rules mechanism

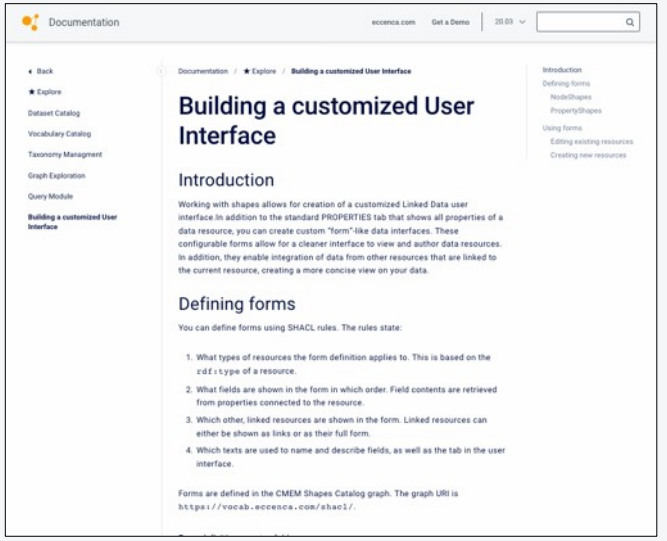




Data defined UI Configuration

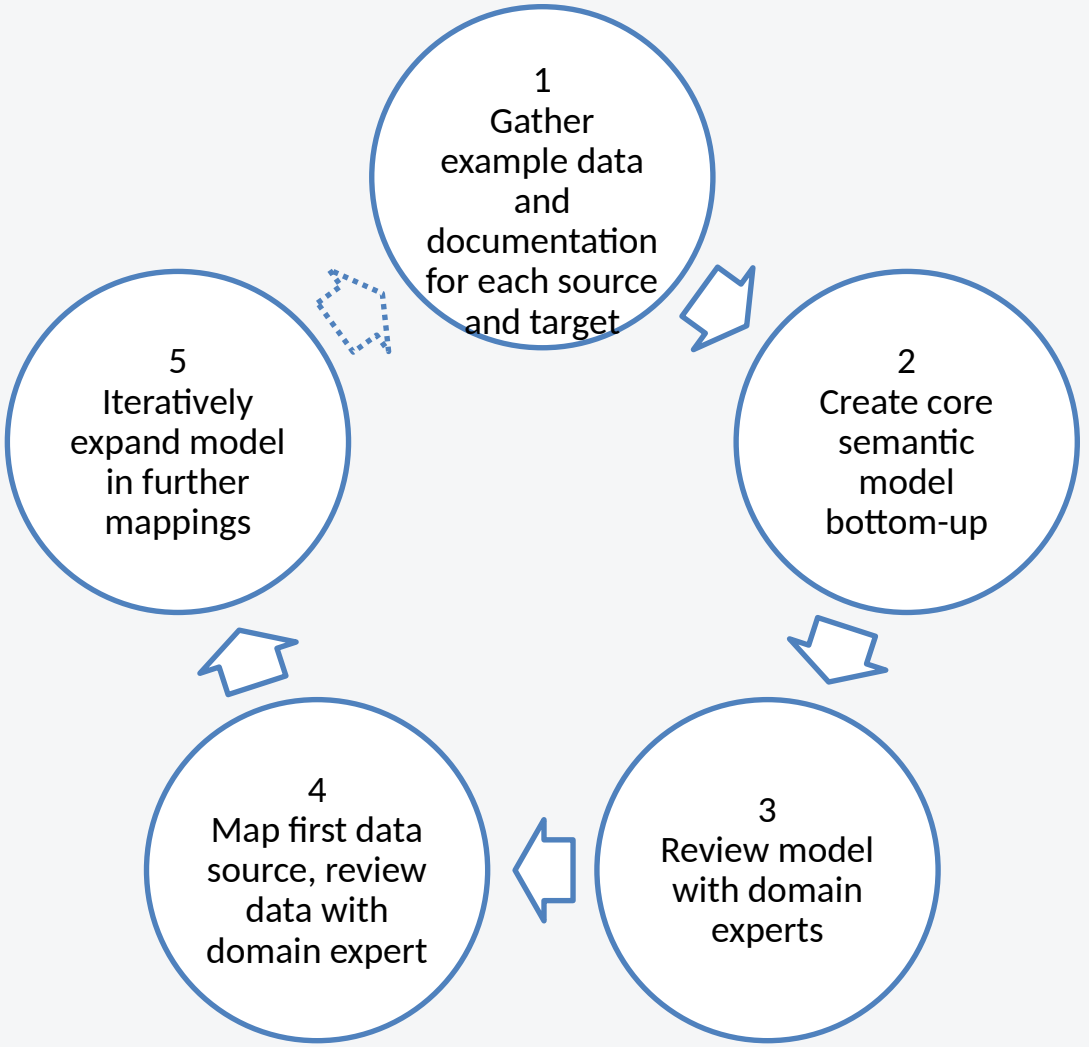
- Framework for interactive data editors that uses W3C SHACL
- Supporting trees of graphs to partition data
- Application UI can be flexibly configured, e.g. including navigation
- Multiple such application configurations can be provided, to support multiple perspectives on the same knowledge graph
- Custom functionality by event driven query hooks





https://documentation.eccenca.com/latest/explore/building-a-customized-user-interface

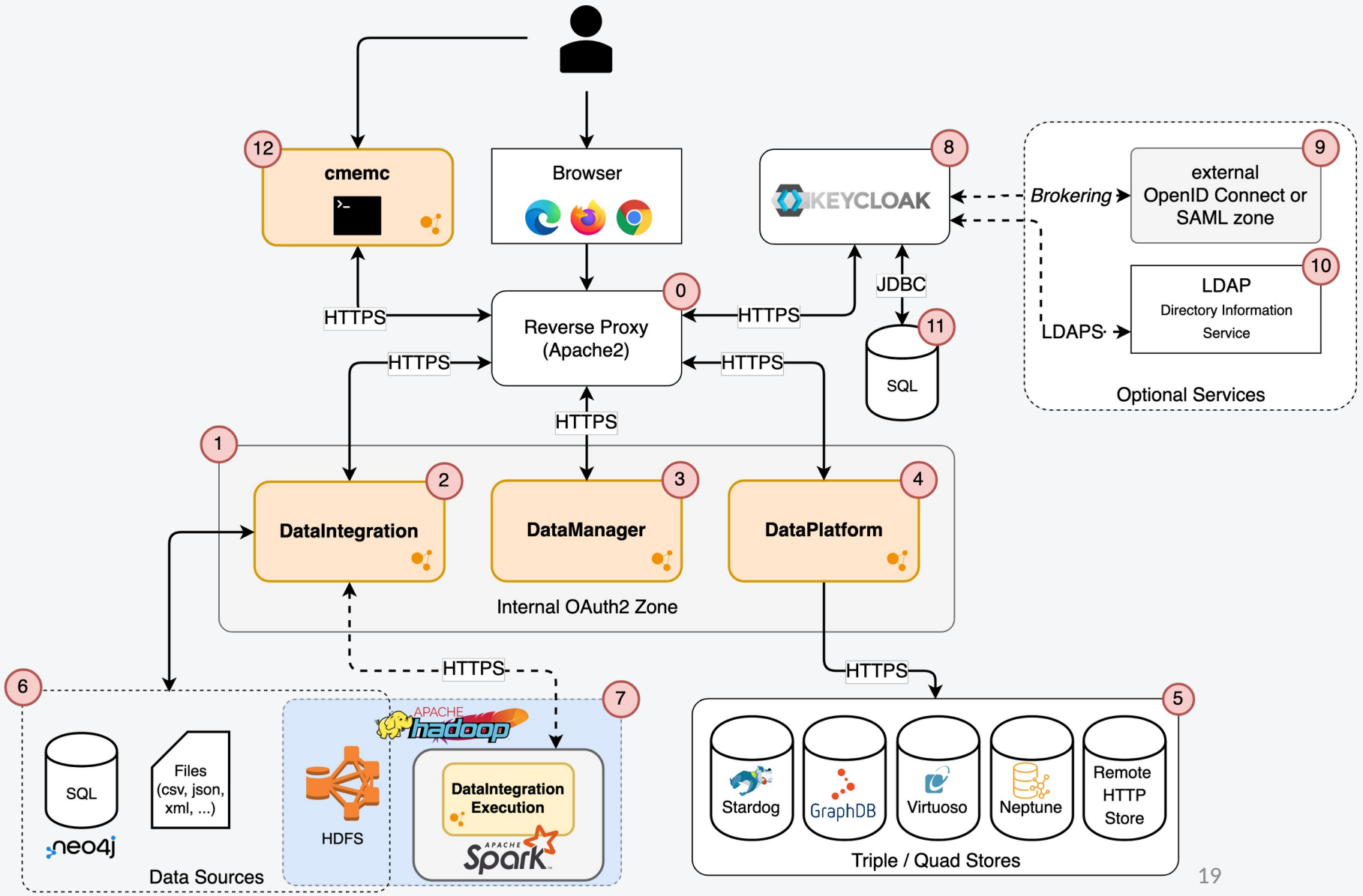
Enterprise Knowledge Graph Project Flow



1. Analyze data sources and understand domain
2. Find or build a semantic model (ontology/vocabulary)
3. Create the mappings
4. Clean, enrich and link data
5. Explore the results

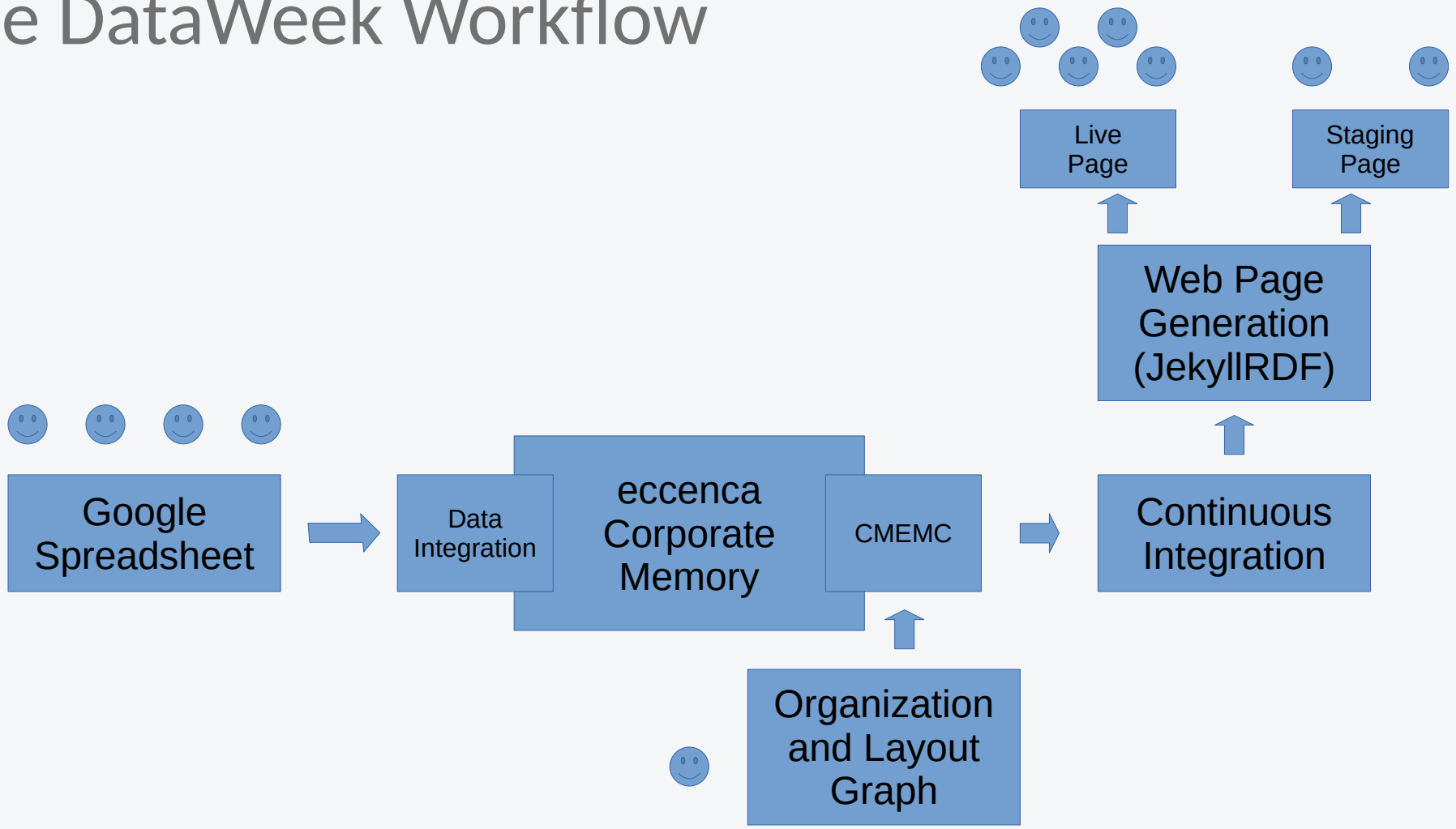
... Re-iterate if needed ...

Reference Architecture Corporate Memory

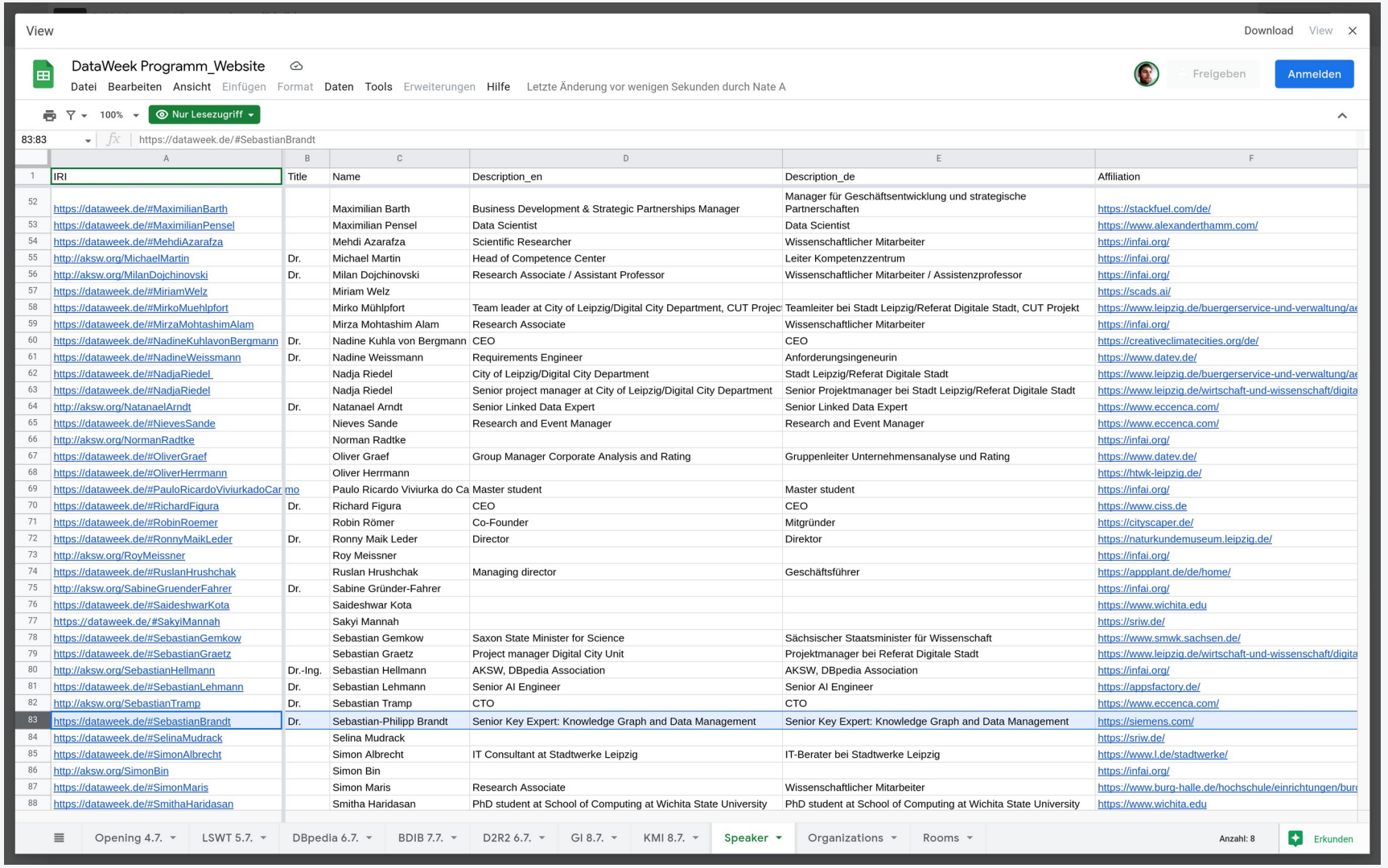


SAMPLE USE CASES

The DataWeek Workflow



The DataWeek Workflow



View — Download — View

DataWeek Programm_Website

Datei Bearbeiten Ansicht Einfügen Format Daten Tools Erweiterungen Hilfe — Letzte Änderung vor wenigen Sekunden durch Nate A

Freigeben — Anmelden

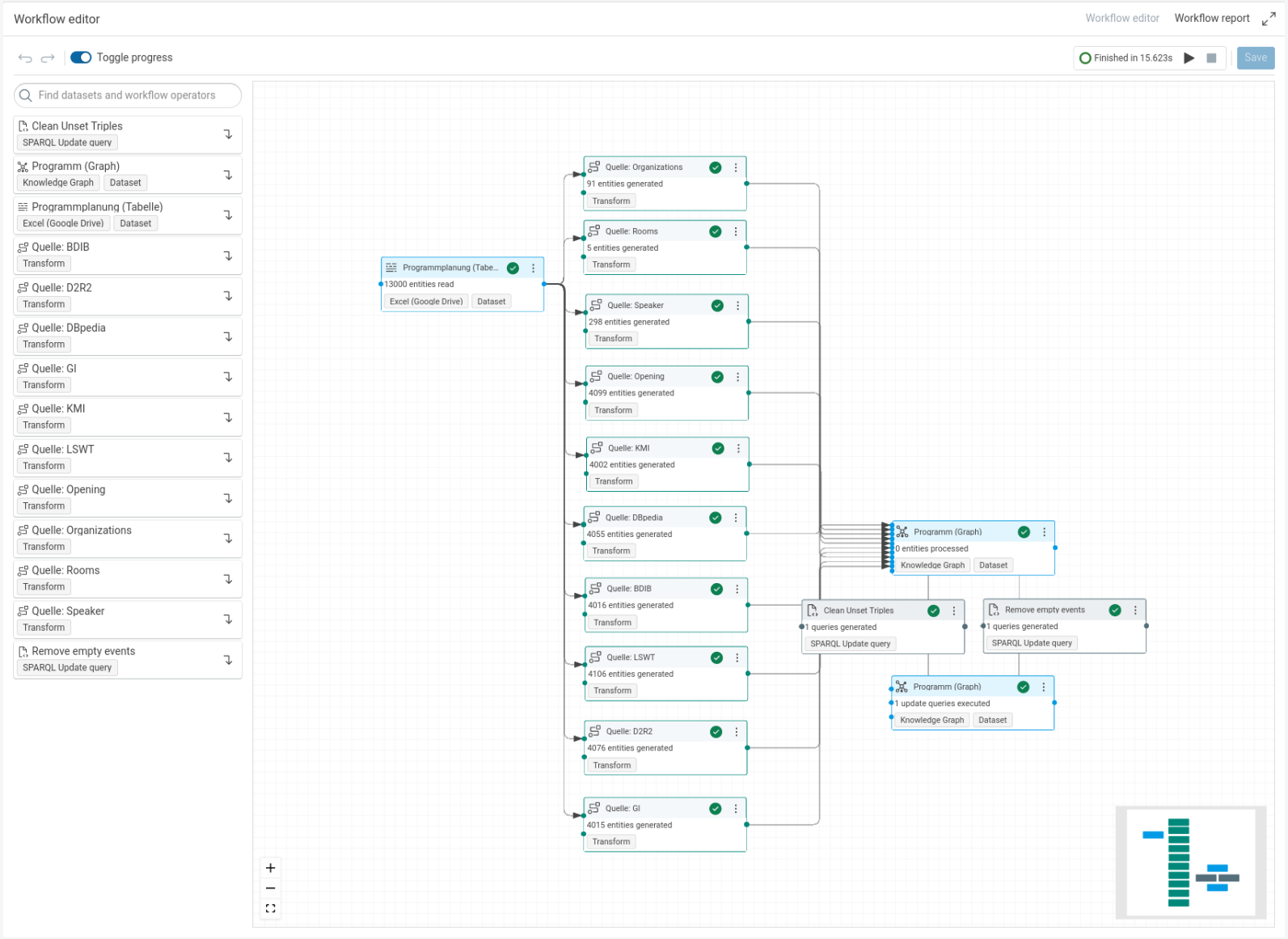
100% — Nur Lesezugriff

83:83 — https://dataweek.de/#SebastianBrandt

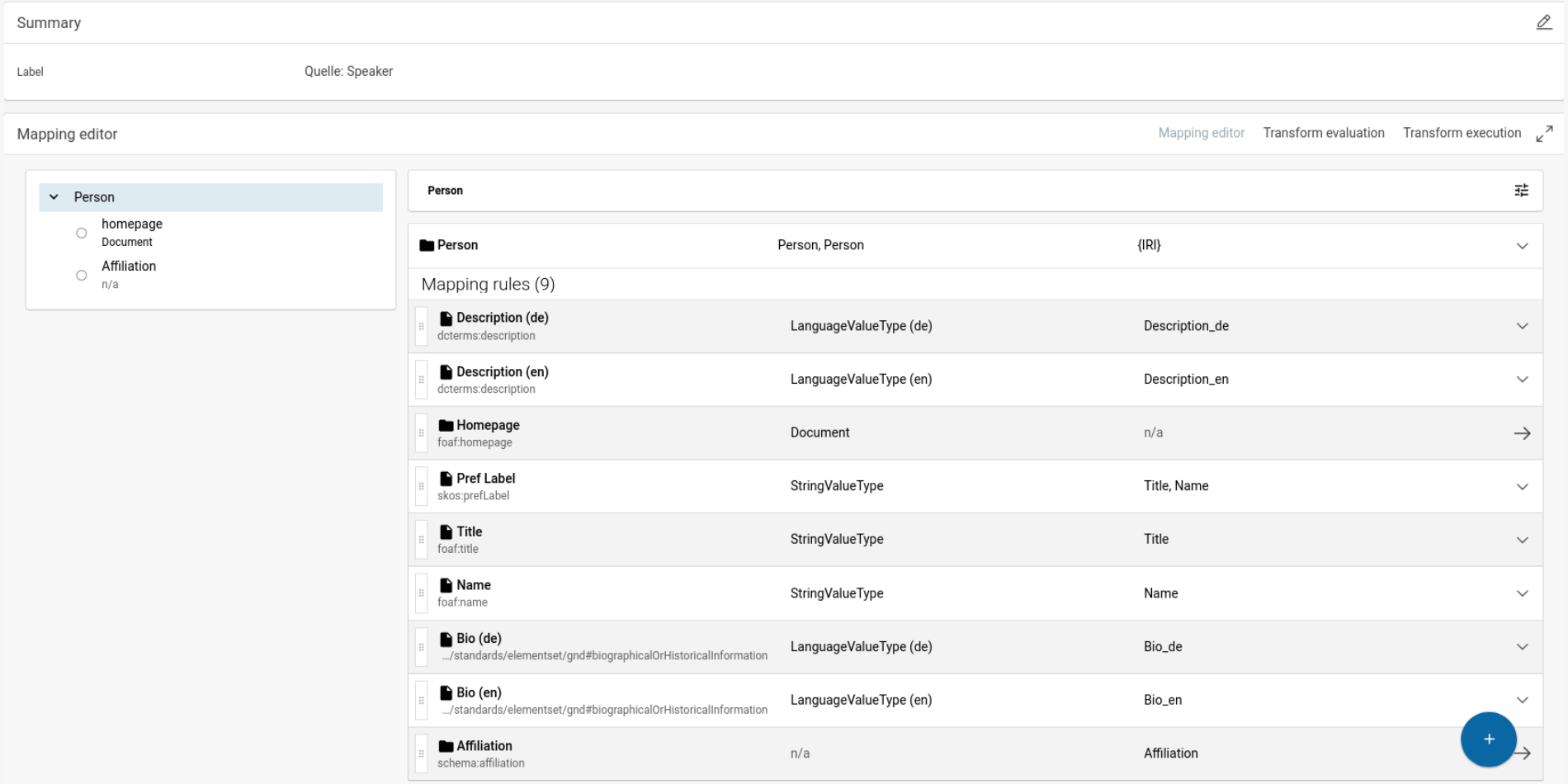
	A	B	C	D	E	F
1	IRI	Title	Name	Description_en	Description_de	Affiliation
52	https://dataweek.de/#MaximilianBarth		Maximilian Barth	Business Development & Strategic Partnerships Manager	Manager für Geschäftsentwicklung und strategische Partnerschaften	https://stackfuel.com/de/
53	https://dataweek.de/#MaximilianPensel		Maximilian Pensel	Data Scientist	Data Scientist	https://www.alexanderthamm.com/
54	https://dataweek.de/#MehdiAzarafza		Mehdi Azarafza	Scientific Researcher	Wissenschaftlicher Mitarbeiter	https://infai.org/
55	http://aksw.org/MichaelMartin	Dr.	Michael Martin	Head of Competence Center	Leiter Kompetenzzentrum	https://infai.org/
56	http://aksw.org/MilanDojchinovski	Dr.	Milan Dojchinovski	Research Associate / Assistant Professor	Wissenschaftlicher Mitarbeiter / Assistenzprofessor	https://infai.org/
57	https://dataweek.de/#MiriamWelz		Miriam Welz			https://scads.ai/
58	https://dataweek.de/#MirkoMuehlpfort		Mirko Mühlpfort	Team leader at City of Leipzig/Digital City Department, CUT Projec	Teamleiter bei Stadt Leipzig/Referat Digitale Stadt, CUT Projekt	https://www.leipzig.de/buergerservice-und-verwaltung/ae
59	https://dataweek.de/#MirzaMohtashimAlam		Mirza Mohtashim Alam	Research Associate	Wissenschaftlicher Mitarbeiter	https://infai.org/
60	https://dataweek.de/#NadineKuhlavonBergmann	Dr.	Nadine Kuhla von Bergmann	CEO	CEO	https://creativeclimatecities.org/de/
61	https://dataweek.de/#NadineWeissmann	Dr.	Nadine Weissmann	Requirements Engineer	Anforderungsingeneurin	https://www.datev.de/
62	https://dataweek.de/#NadjaRiedel		Nadja Riedel	City of Leipzig/Digital City Department	Stadt Leipzig/Referat Digitale Stadt	https://www.leipzig.de/buergerservice-und-verwaltung/ae
63	https://dataweek.de/#NadjaRiedel		Nadja Riedel	Senior project manager at City of Leipzig/Digital City Department	Senior Projektmanager bei Stadt Leipzig/Referat Digitale Stadt	https://www.leipzig.de/wirtschaft-und-wissenschaft/digita
64	http://aksw.org/NatanaelArndt	Dr.	Natanael Arndt	Senior Linked Data Expert	Senior Linked Data Expert	https://www.eccenca.com/
65	https://dataweek.de/#NievesSande		Nieves Sande	Research and Event Manager	Research and Event Manager	https://www.eccenca.com/
66	http://aksw.org/NormanRadtke		Norman Radtke			https://infai.org/
67	https://dataweek.de/#OliverGraef		Oliver Graef	Group Manager Corporate Analysis and Rating	Gruppenleiter Unternehmensanalyse und Rating	https://www.datev.de/
68	https://dataweek.de/#OliverHerrmann		Oliver Herrmann			https://htwk-leipzig.de/
69	https://dataweek.de/#PauloRicardoViviurkadoCar	mo	Paulo Ricardo Viviurka do Ca	Master student	Master student	https://infai.org/
70	https://dataweek.de/#RichardFigura	Dr.	Richard Figura	CEO	CEO	https://www.ciss.de
71	https://dataweek.de/#RobinRoemer		Robin Römer	Co-Founder	Mitgründer	https://cityscaper.de/
72	https://dataweek.de/#RonnyMaikLeder	Dr.	Ronny Maik Leder	Director	Direktor	https://naturkundemuseum.leipzig.de/
73	http://aksw.org/RoyMeissner		Roy Meissner			https://infai.org/
74	https://dataweek.de/#RuslanHrushchak		Ruslan Hrushchak	Managing director	Geschäftsführer	https://appplant.de/de/home/
75	http://aksw.org/SabineGruenderFahrer	Dr.	Sabine Gründer-Fahrer			https://infai.org/
76	https://dataweek.de/#SaideshwarKota		Saideshwar Kota			https://www.wichita.edu
77	https://dataweek.de/#SakyiMannah		Sakyi Mannah			https://sriw.de/
78	https://dataweek.de/#SebastianGemkow		Sebastian Gemkow	Saxon State Minister for Science	Sächsischer Staatsminister für Wissenschaft	https://www.smwk.sachsen.de/
79	https://dataweek.de/#SebastianGraetz		Sebastian Graetz	Project manager Digital City Unit	Projektmanager bei Referat Digitale Stadt	https://www.leipzig.de/wirtschaft-und-wissenschaft/digita
80	http://aksw.org/SebastianHellmann	Dr.-Ing.	Sebastian Hellmann	AKSW, DBpedia Association	AKSW, DBpedia Association	https://infai.org/
81	https://dataweek.de/#SebastianLehmann	Dr.	Sebastian Lehmann	Senior AI Engineer	Senior AI Engineer	https://appsfactory.de/
82	http://aksw.org/SebastianTramp	Dr.	Sebastian Tramp	CTO	CTO	https://www.eccenca.com/
83	https://dataweek.de/#SebastianBrandt	Dr.	Sebastian-Philipp Brandt	Senior Key Expert: Knowledge Graph and Data Management	Senior Key Expert: Knowledge Graph and Data Management	https://siemens.com/
84	https://dataweek.de/#SelinaMudrack		Selina Mudrack			https://sriw.de/
85	https://dataweek.de/#SimonAlbrecht		Simon Albrecht	IT Consultant at Stadtwerke Leipzig	IT-Berater bei Stadtwerke Leipzig	https://www.l.de/stadtwerke/
86	http://aksw.org/SimonBin		Simon Bin			https://infai.org/
87	https://dataweek.de/#SimonMaris		Simon Maris	Research Associate	Wissenschaftlicher Mitarbeiter	https://www.burg-halle.de/hochschule/einrichtungen/bur
88	https://dataweek.de/#SmithaHaridasan		Smitha Haridasan	PhD student at School of Computing at Wichita State University	PhD student at School of Computing at Wichita State University	https://www.wichita.edu

Opening 4.7. — LSWT 5.7. — DBpedia 6.7. — BDIB 7.7. — D2R2 6.7. — GI 8.7. — KMI 8.7. — Speaker — Organizations — Rooms — Anzahl: 8 — Erkunden

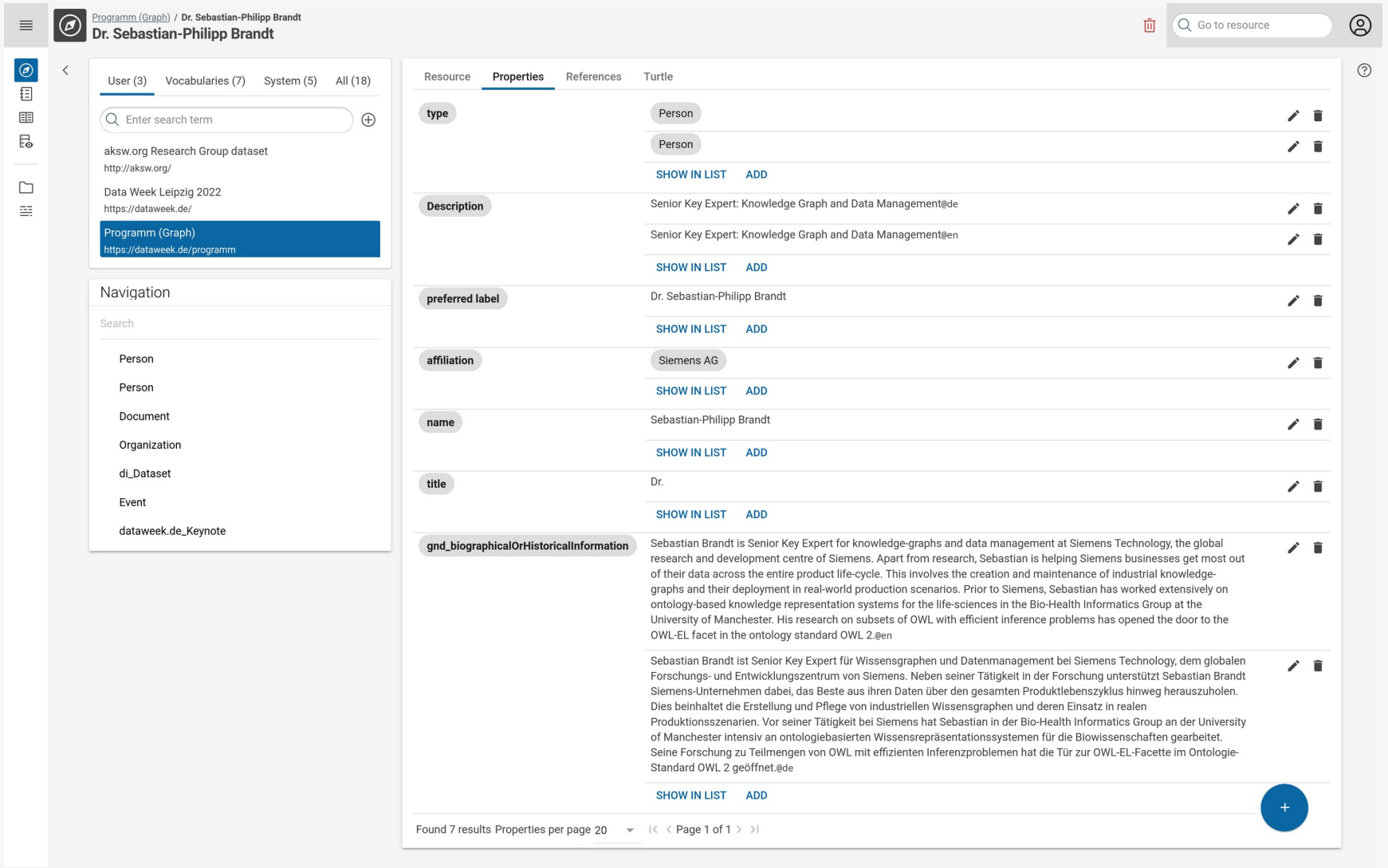
The DataWeek Workflow



The DataWeek Workflow



The DataWeek Workflow



eccenca GmbH
Hainstraße 8
D-04109 Leipzig
Germany

+49 341 2650 8028
info@eccenca.com
https://eccenca.com

eccenca
mastering complexity

Hans-Chr. Brockmann
Geschäftsführer

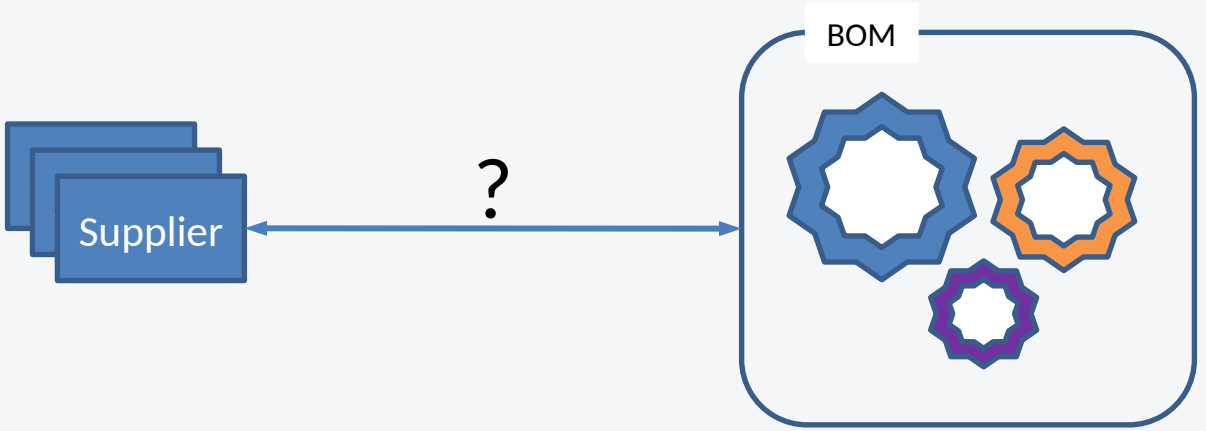
Hainstraße 8
D-04109 Leipzig
Germany

+49 511 3365 2810
+49 173 3698 610

brockmann@eccenca.com
https://eccenca.com

BACKUP

Apply Machine Learning to perform “Form Fit Functional” Material Linking



Today based on part number matching and man-made part lists:

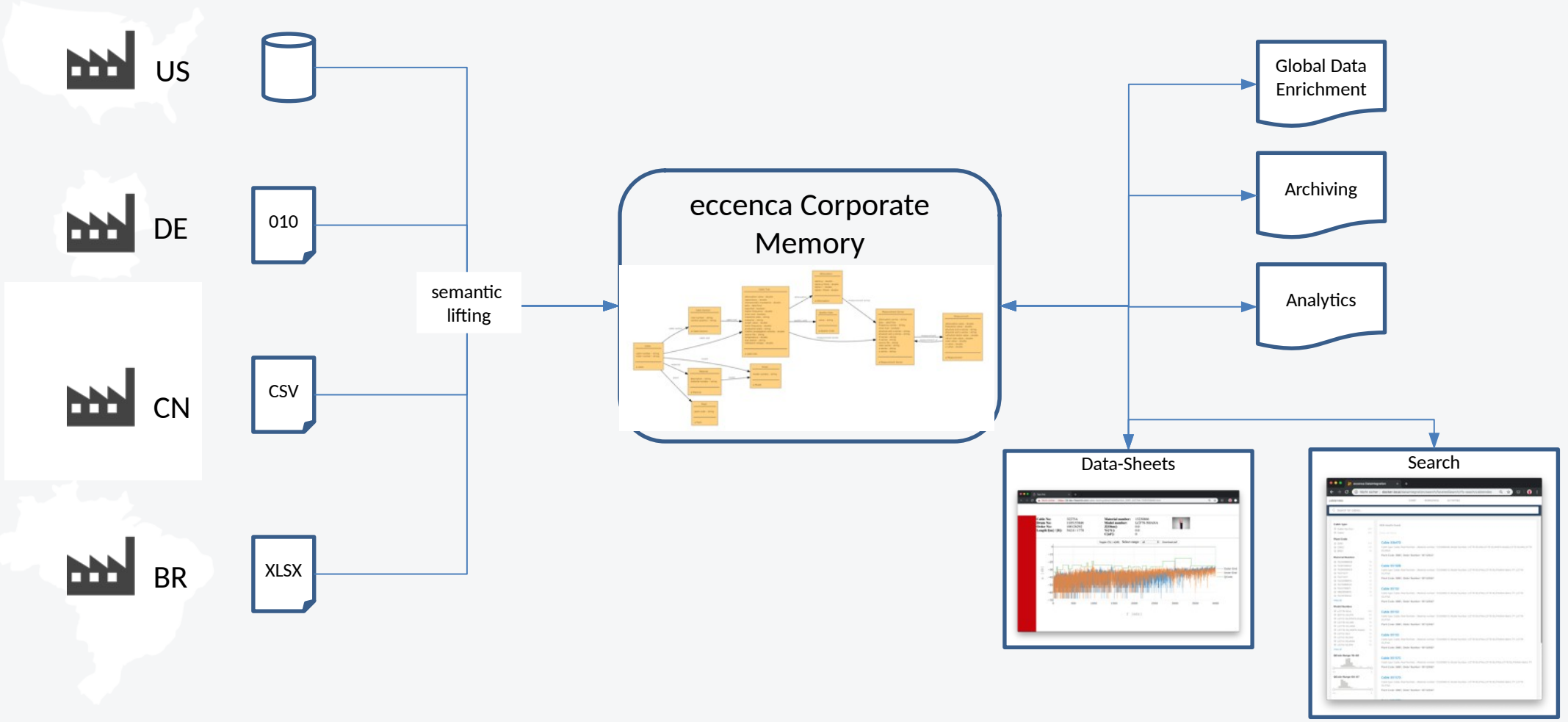
MPN	**Supplier**
293D105X9016A2##E3	VISHAY
B45196E3105K10	KEMET
T491A105K016AS	KEMET
T491A105K016AT	KEMET
TAJA105K016RNJ	AVX

Machine learning based on granular material properties like:

Capacitor Type	TANTALUM CAPACITOR
Capacitance	1.0 µF
Dielectric Material	TANTALUM (DRY/SOLID)
Mounting Feature	SURFACE MOUNT
Neg. Tolerance	10.0 %
Op. Temp.-Min	-55.0 Cel
Op. Temp.-Max	85.0 Cel
Package Shape	RECTANGULAR PACKAGE
Package Style	SMT

…

Cross Site Semantic Data Harmonization



Project Benefits

- 12% Inventory reduction contribution within 3 months
- 200% Project ROI Data-as-Service creates automation and productivity gains
- 70% of effort analyzing data
  Instead of finding, aligning, integrating and cleaning data

About eccenca GmbH

brox IT-Solutions GmbH

brox information excellence

Gegründet:	1998
Fokus:	IT-Consulting
IPR:	Initiierung/Leitung eclipse.org/SMILA eccenca Enterprise Search
Key Accounts:	Volkswagen, Audi, Skoda, MAN, Telekom, Daimler, Bosch, Siemens, Continental

AKSW – Universität Leipzig & Fraunhofer IAIS

AKSW

Führende Linked Data Forschungseinheit in Europa. Initiator des nationalen Industrial Data Space

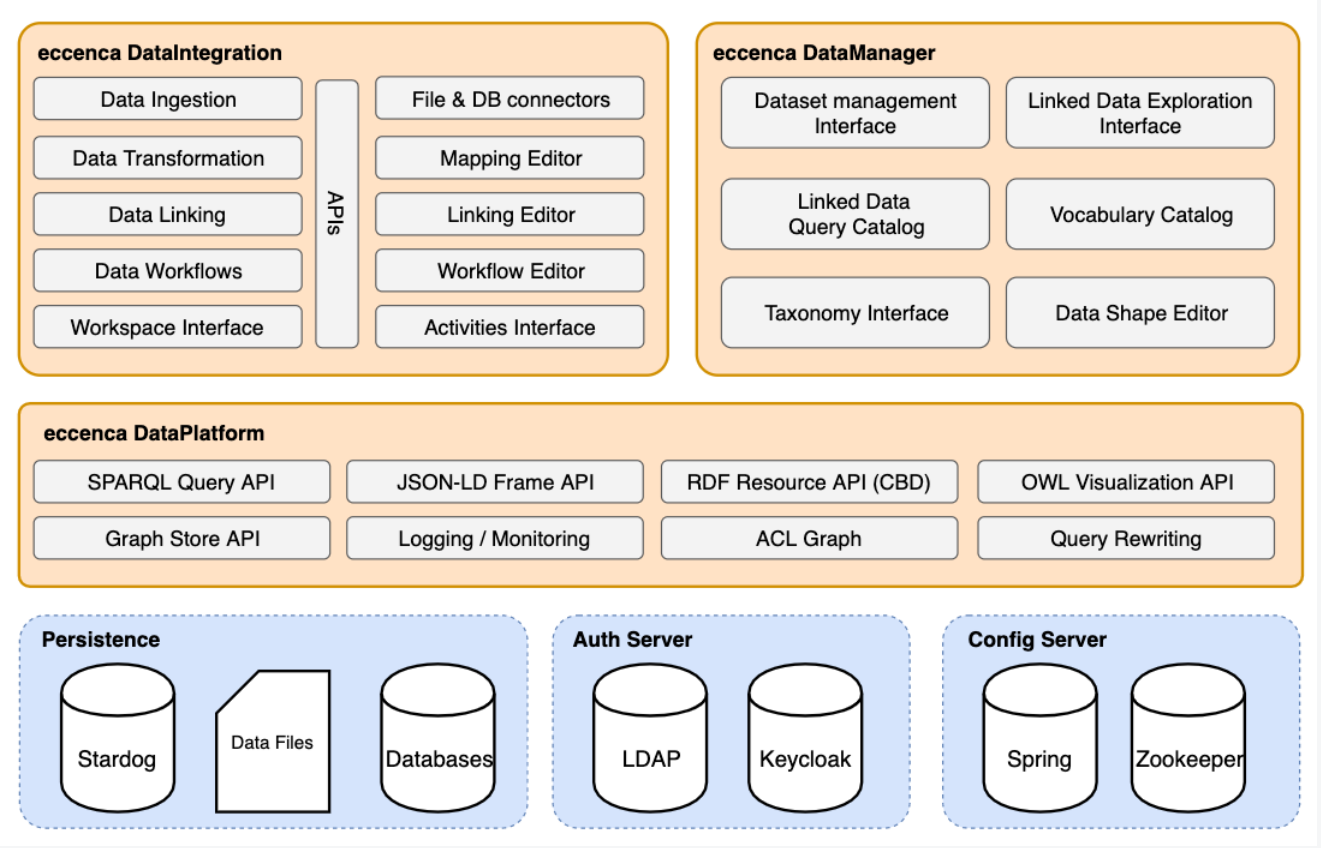
Initiator:	DBPedia, Linked GeoData etc.
Betreiber:	Datenportal der EU-Kommission
Team-Leitung:	Prof. Dr. Sören Auer

eccenca GmbH

eccenca Command your Data!

Gründung:	2013
Fokus:	Produkte/Lösungen
Team Size:	35
IPR:	Linked Data, M2N Synchronization, Linking, Authentication/Data Security
Kunden:	Volkswagen, Bosch, Nokia, Infineon, Ericsson, Telekom, Daimler
F&E Projekte:	LUCID, ELDS, GeoKnow, Diachron
Initiativen:	MOBIVOC, OSFP

Technical Architecture



eccenca DataManager – Features

- Management of Knowledge Bases (Named Graphs, Linked Data access optional)
- Tree, list and resource views
- Versioning (triple based)
- User management and access control
- Query Catalog
- Dataset Schema Browser

- Inline authoring
- Detailed edit view
- Add new resources and properties
- Search
- Facet based filtering
- Complex navigation hierarchies (spatial, class based, organization structure based)
-

eccenca DataIntegration – Features

- UI allows to view and edit linkage rules
- Linkage rules are shown as a tree
- Editing using drag & drop
- DataIntegration provides a high level data manipulation and linking engine
  - Execution of linkage and integration rules on arbitrary datasets provided by eccenca DataPlatform
  - Manual creation of rules with an integrated editor
  - Automatic learning of rules based on training data (positive / negative lists)
-

eccenca DataPlatform – Features

- DataPlatform is a semantic middleware which provides a unified access to structured data
  - W3C standards such as RDF / Linked Data and SPARQL
  - Authorization based on an RDF Access Control Model
  - Authentication via OAuth2 protocol
  - Integration with external User Management Systems (e.g. LDAP, Active Directory)
  - Versioning Support (tracking of changes on triples and graphs)
  - Integration of non-RDF sources via mapping and query rewriting technologies (RDB2RDF component)

Semantic Data Management
Data in Context

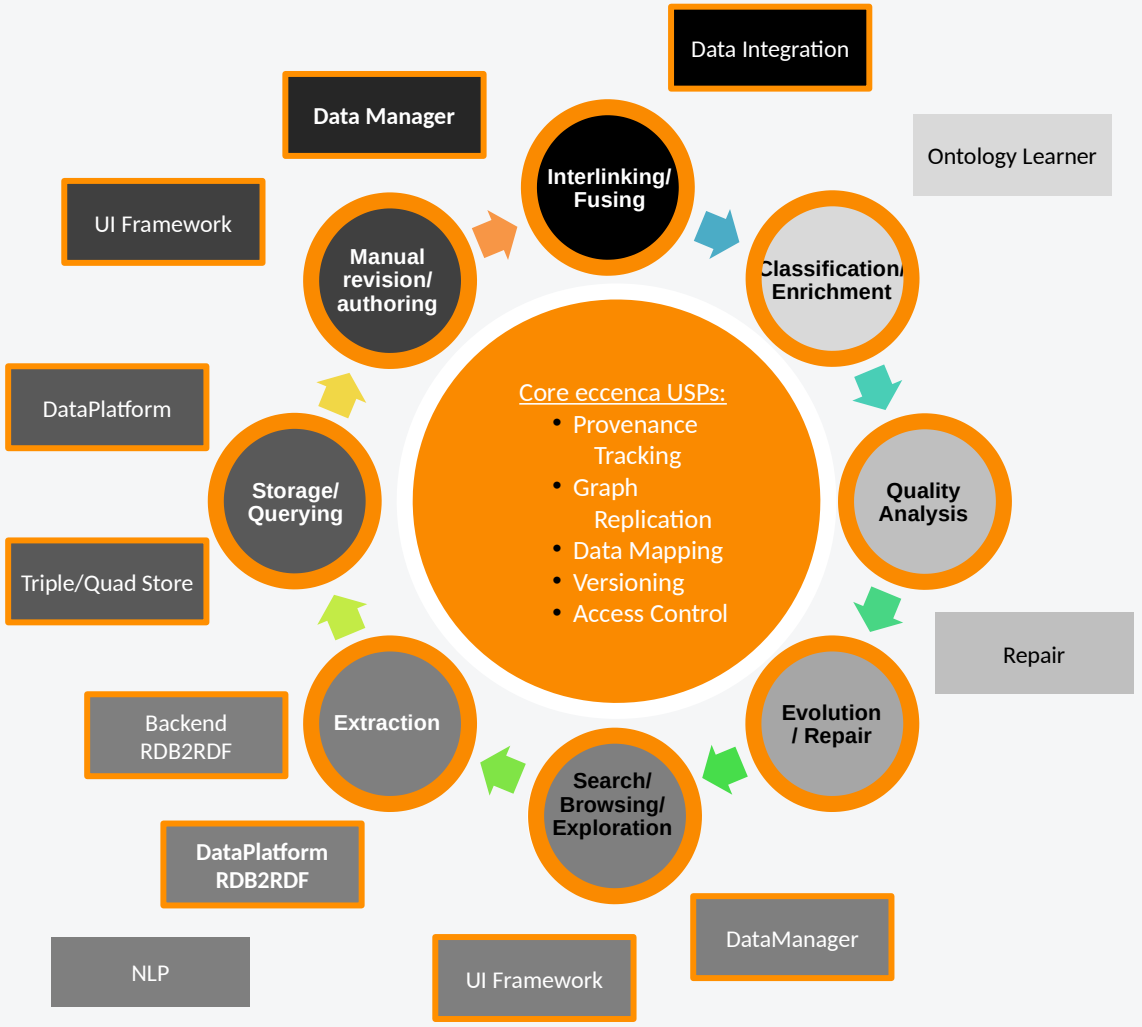
What

- Create knowledge graphs by connecting datasets and metadata to logical models
- Physical data models unchanged!
-
  - Explore metadata & structures
  - Query & Access data via models
  - Integrate data on model level
  - Share data on model level

How

- Leverage linked data principles
  - Schema as data (RDF)
  - Global identifiers (URIs)
  - Linked data graphs (W3C)
-
- Catalog your data assets: datasets, vocabularies (models), …
- Publish-subscribe for sharing
- Machine learning for integration

Linked Data Life Cycle



- Extraction / Mapping
- Storage / Querying
- Manual Revision / Authoring
- Linking / Fusion
- Classification / Enrichment
- Quality / Evolution
- Search / Browse / Explore

Changing the Data-Collaboration Paradigm

… by turning STRINGS into THINGS

Schema defines Semantics

API
EDI
XML

Local Standardization
Slow Innovation
NxM cost
Changes break the contract

RDF

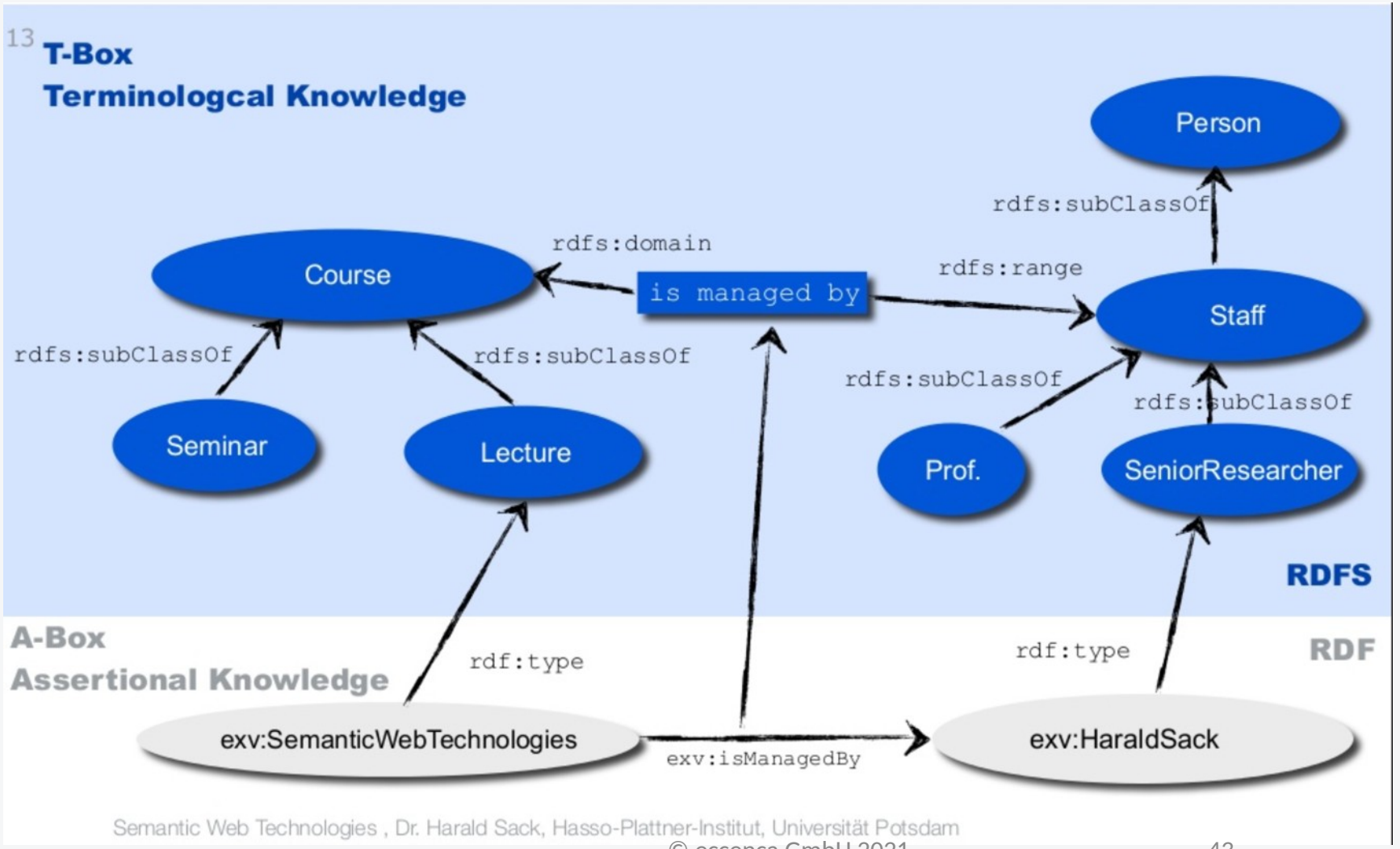
Standardized Schema (SPO)
Vocabularies define Semantics

Low Standardization
Agile Innovation
N+1 cost
Changes amend the contract

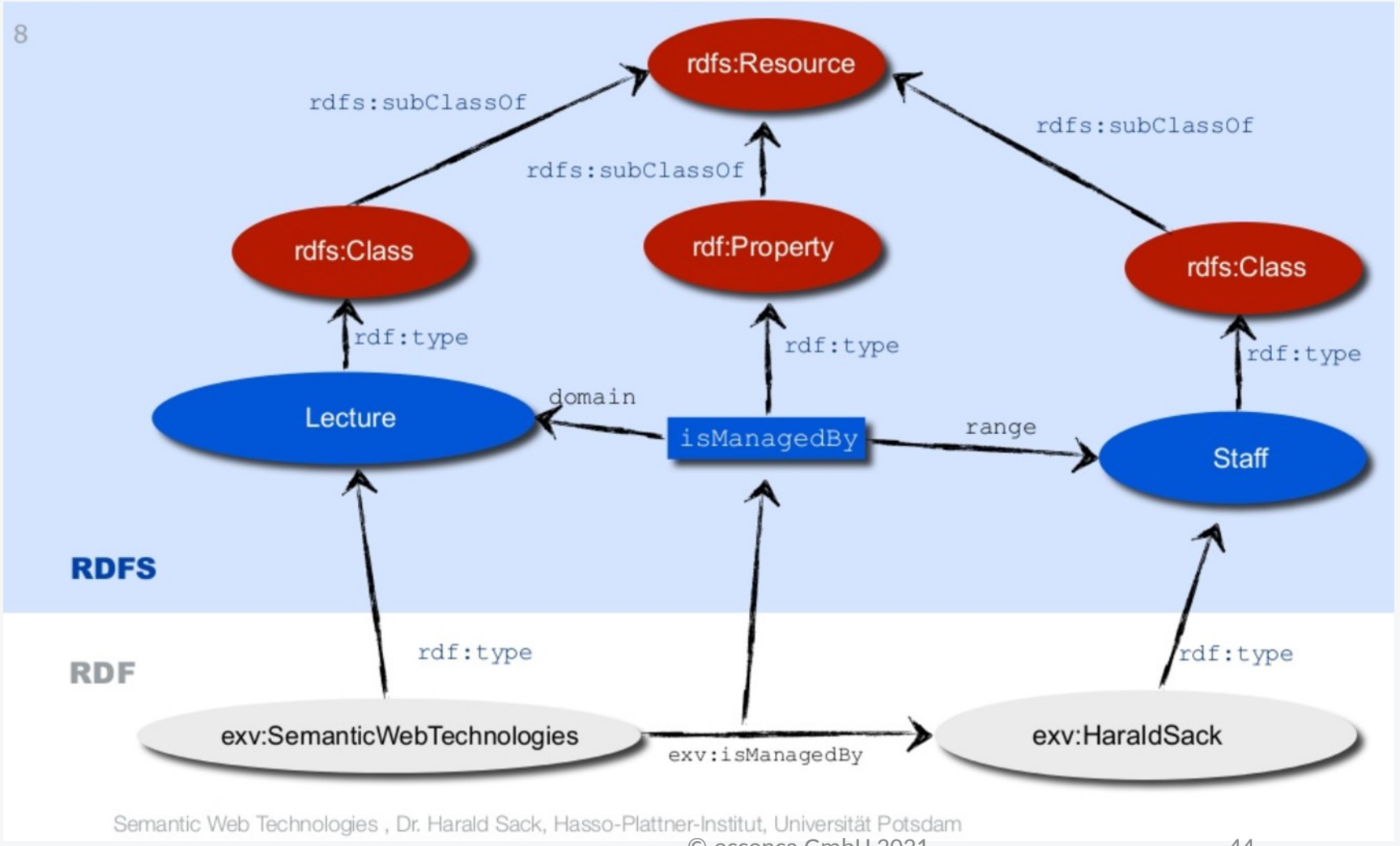
Other Use Cases

- Conceptual / semantic data model based data preparation for BI and analytics
- Yield management / lead time based dynamic pricing
- Data integration and central data hub for the Software Monetarization Platform project
- Semantic Enterprise Information Model
  - MDM, simplification, data preparation, analytics, etc.

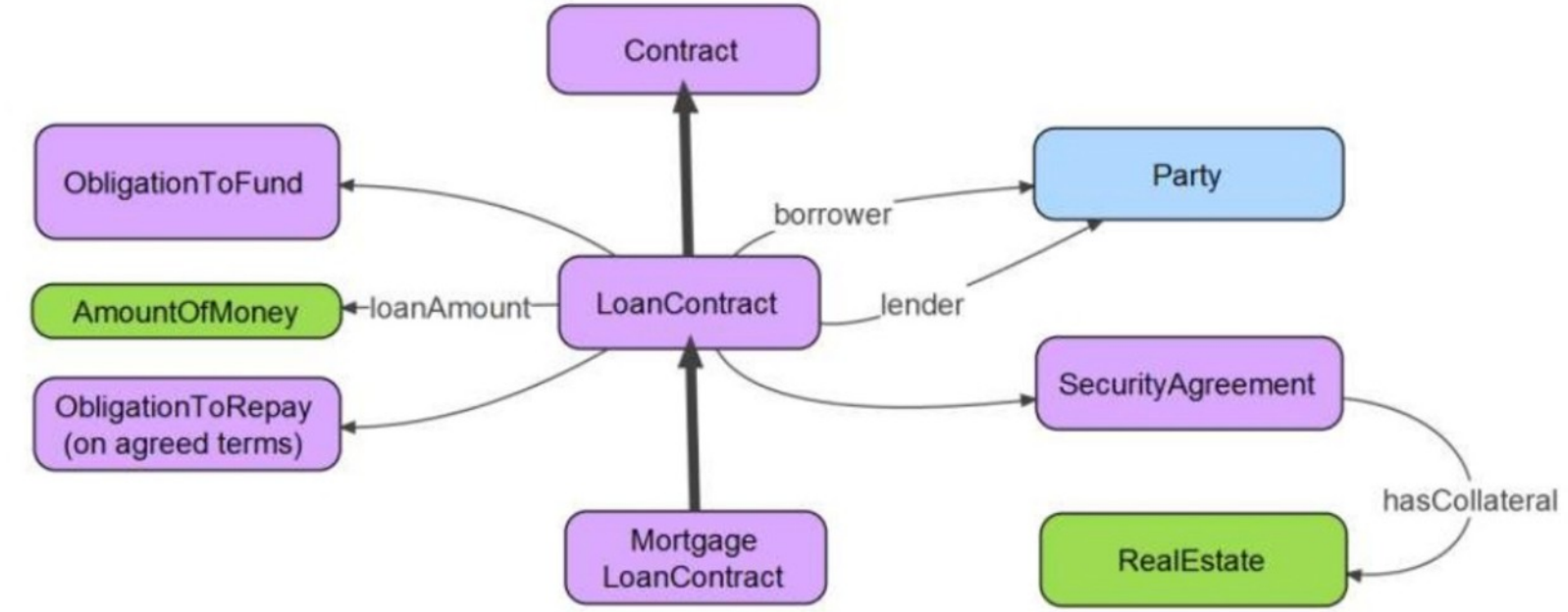
RDF +RDFS



Classes, Properties, Instances



OWL: Inferencing classification.



Mortgage: A LoanContract that has a SecurityAgreement where the collateral is RealEstate. Can infer into this class.

SHACL: Checking Graph Patterns

Constraints on values with another shape

Constraint	Description
node*	All values of a given property must have a given shape Recursion is not allowed in current SHACL

```
:User a sh:NodeShape, rdfs:Class ;
  sh:property [
   sh:path schema:worksFor ;
   sh:node :Company ;
  ] .

:Company a sh:Shape ;
  sh:property [
   sh:path     schema:name ;
   sh:datatype xsd:string ;
  ] .
```

```
:alice a :User;
       schema:worksFor :OurCompany .

:bob   a :User;
       schema:worksFor :Another .

:OurCompany
    schema:name "OurCompany" .

:Another
    schema:name 23 .
```

☹