



Building and exploring Knowledge Graphs with eccenca Corporate Memory

Dr. Natanael Arndt
Senior Linked Data Expert @ eccenca GmbH

Leipzig ASCM - APICS
Business Digital Twin BOSCO
EDM Daimler BERLIN
Council Innovate with Data
>100% Growth NOKIA
Volkswagen 60 FTE
Knowledge Graph VOITH
DBpedia DaaS/Cloud

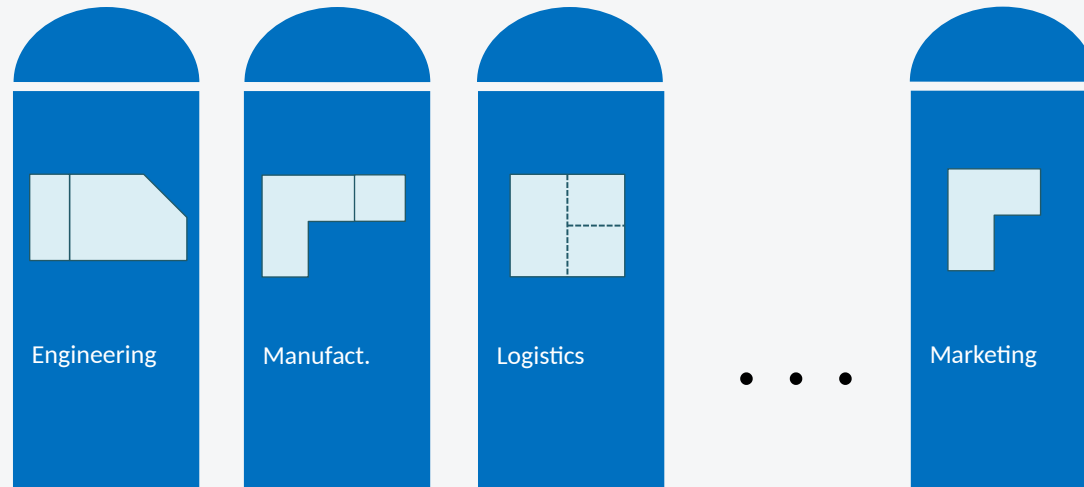
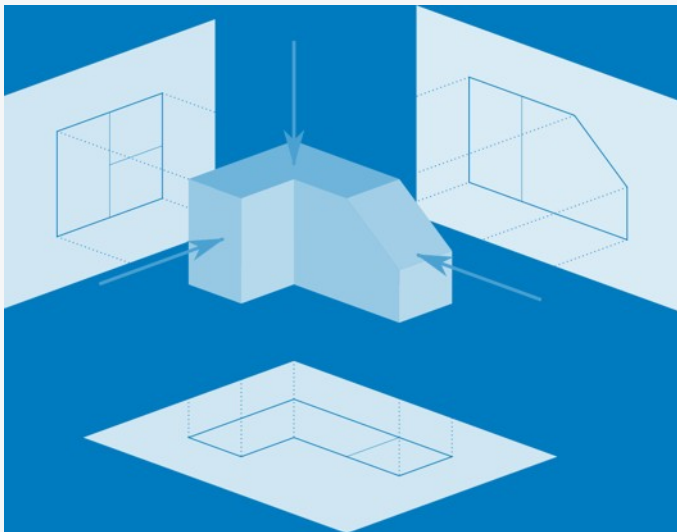


Building and exploring Knowledge Graphs with

eccenca Corporate Memory

Perspectives on data turn into silos

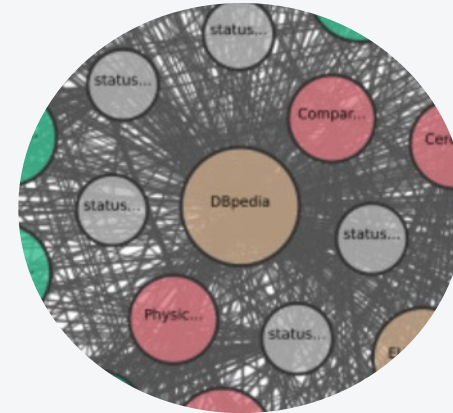
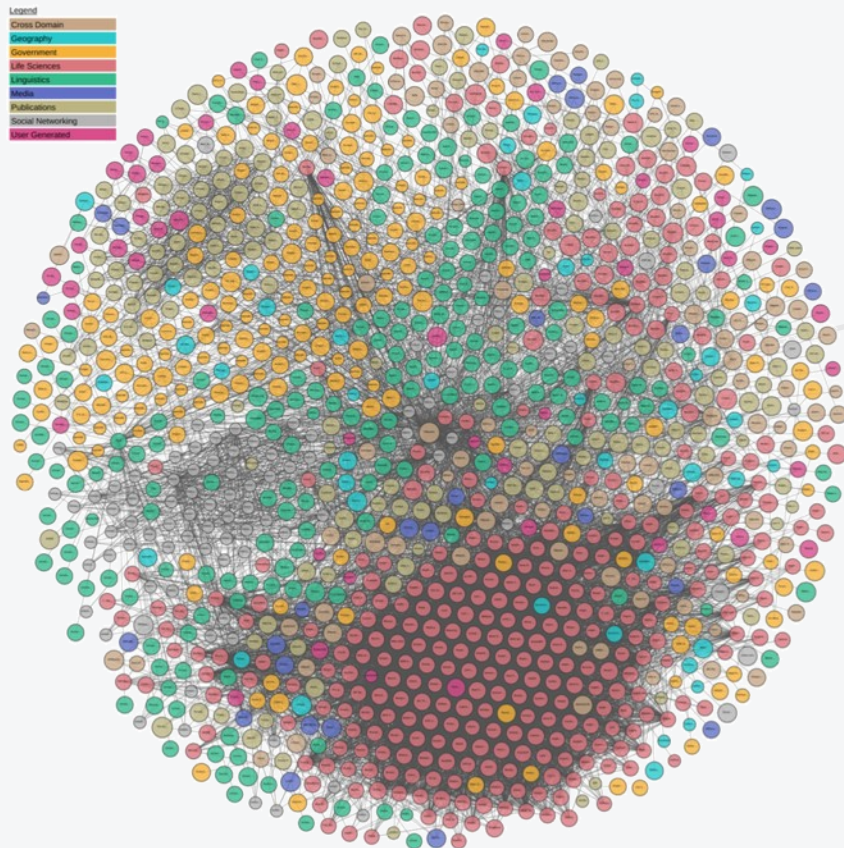
Parts of data are being curated, duplicated, annotated and simply changed over time, making reconciliation and interpretation a challenge



Data on the web & WikiData & DBpedia

Global and Unified Access to Knowledge

The **Linked Open Data** Cloud



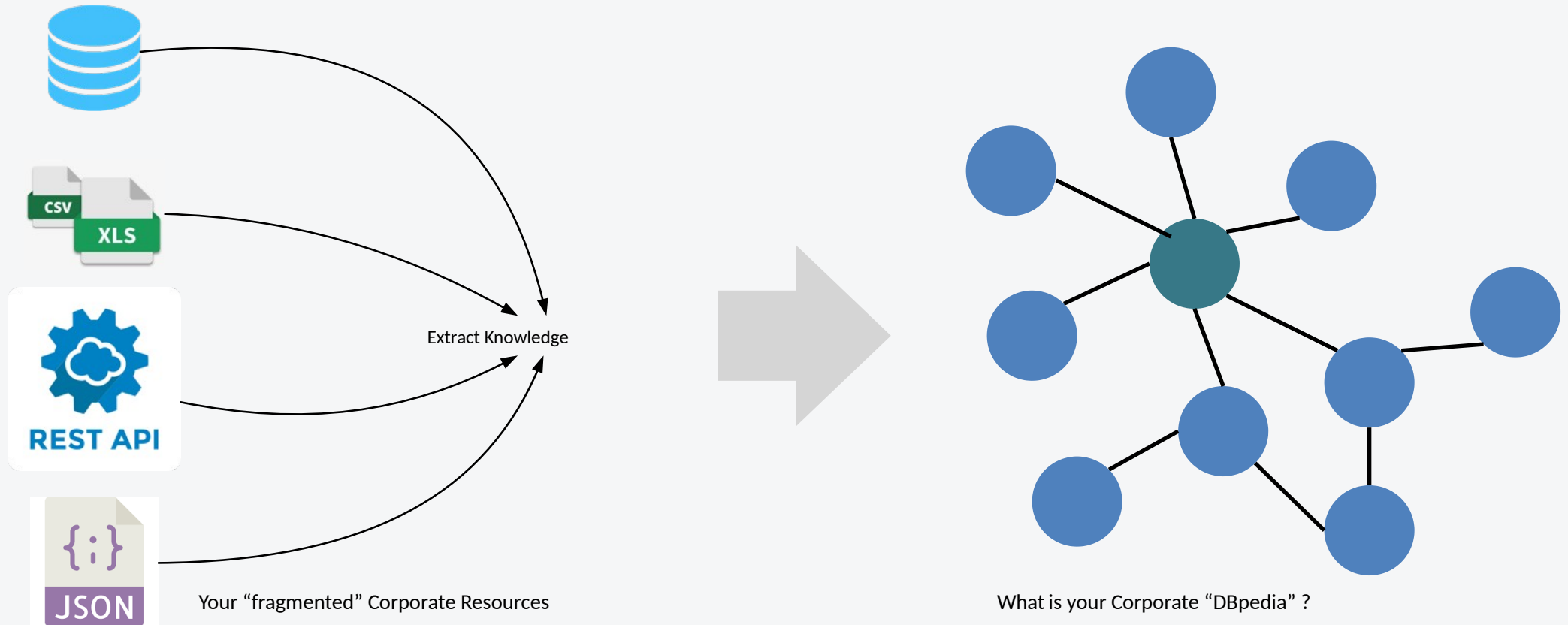
DBpedia extracts structured content from Wikipedia.

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

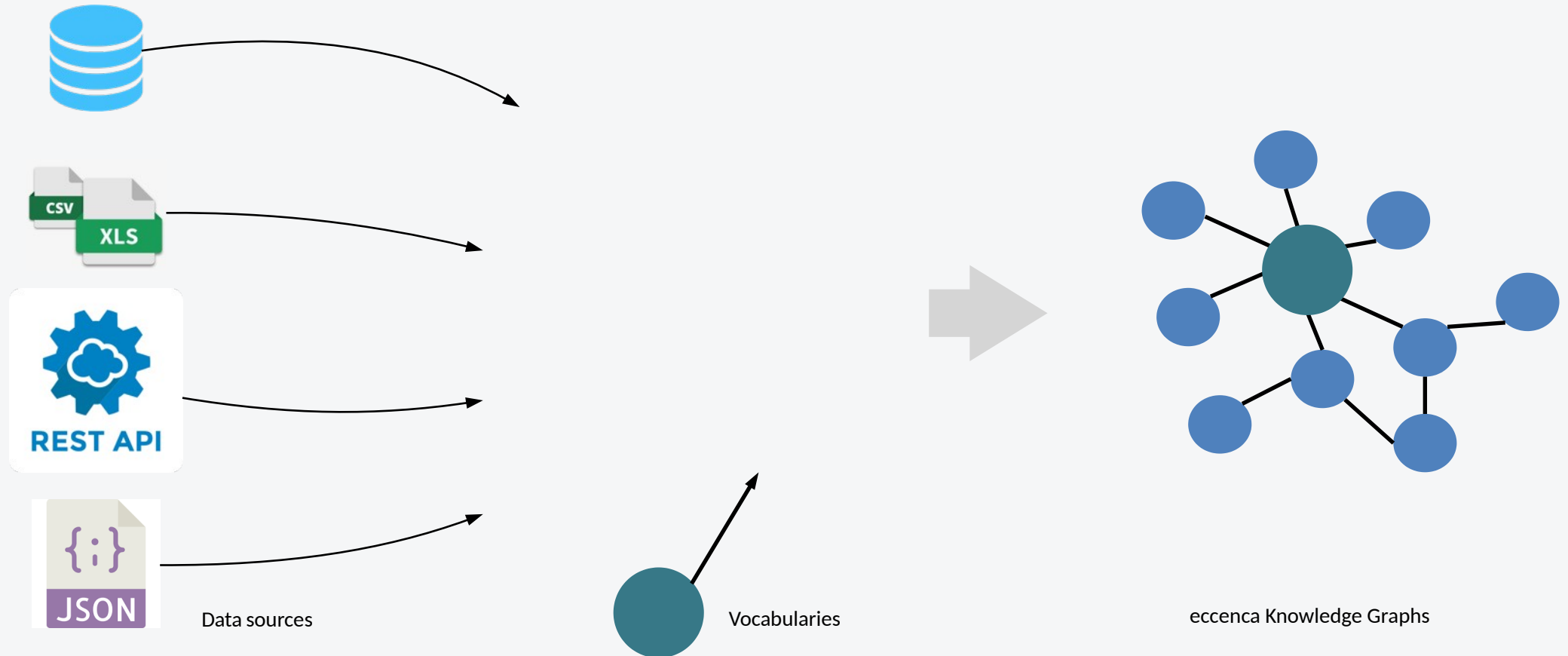
Leipzig University - Prof. Dr. Sören Auer, co-founder of Eccenca

From the web to linked enterprise data

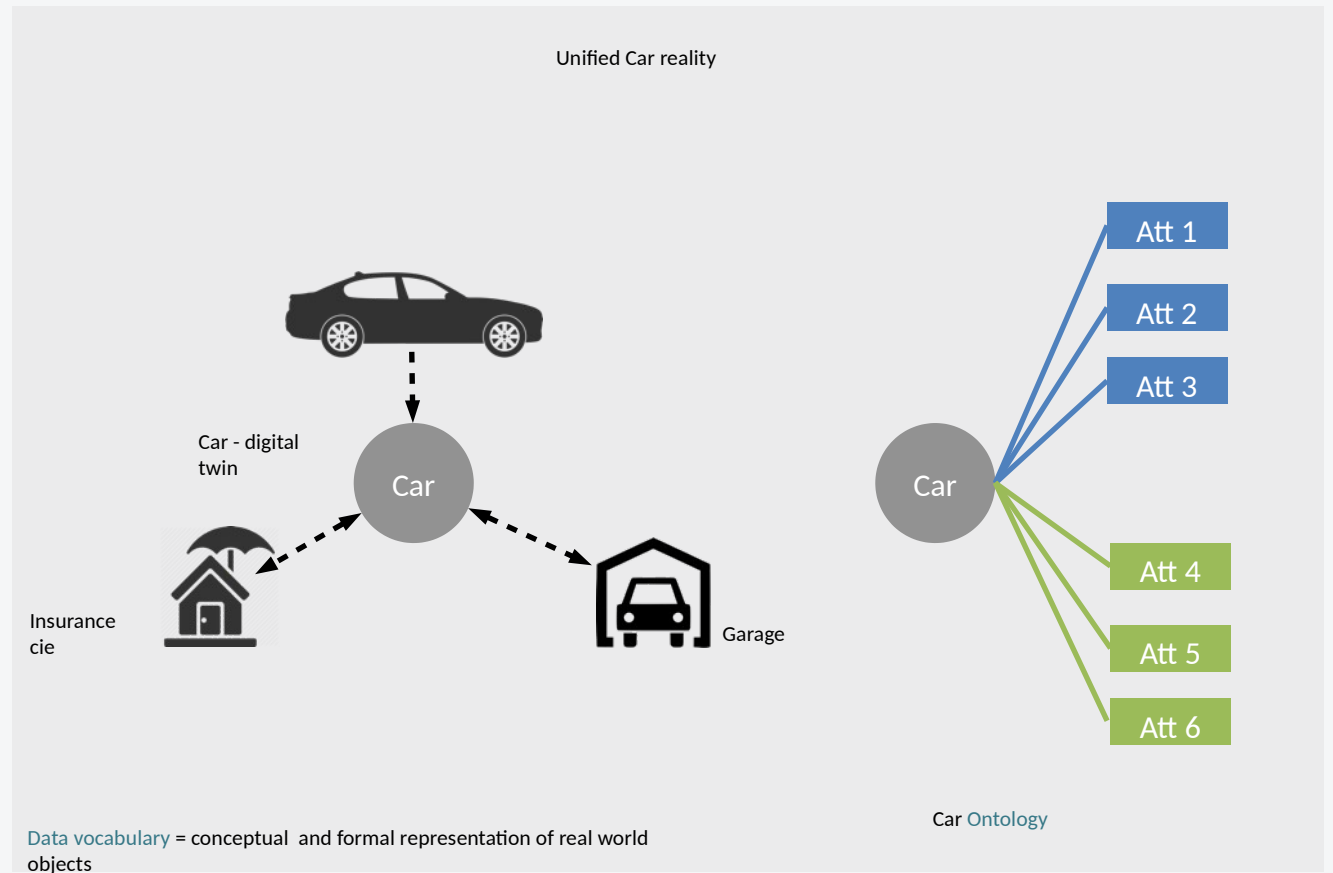
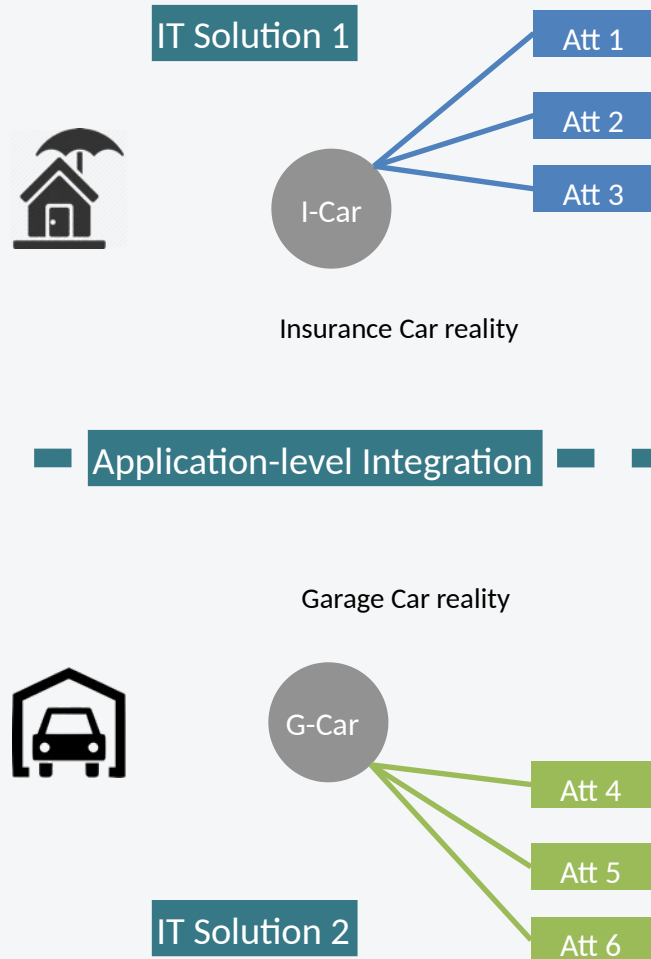
Global & unified access to corporate knowledge



eccenca Corporate Memory

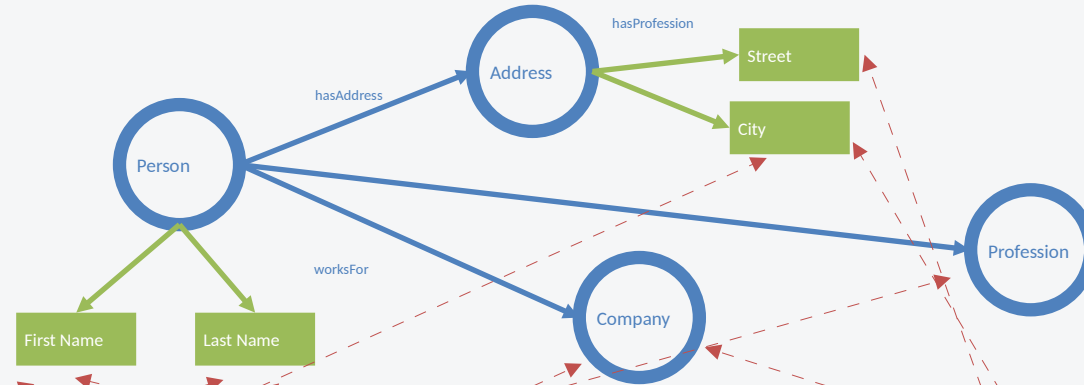


Global identification, multiple perspectives



Semantic metadata vs flat metadata

Rich Semantic Metadata



Automated Validation of Consistency...

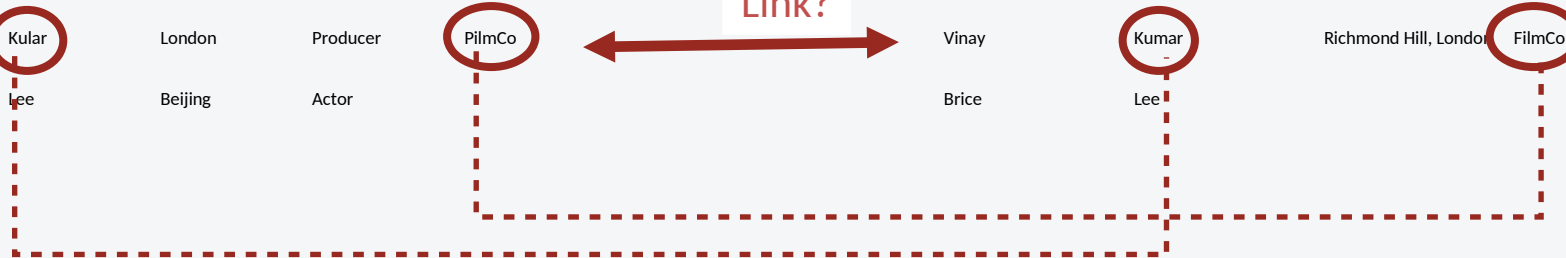
(Semi-)Automated Mapping

Flat metadata

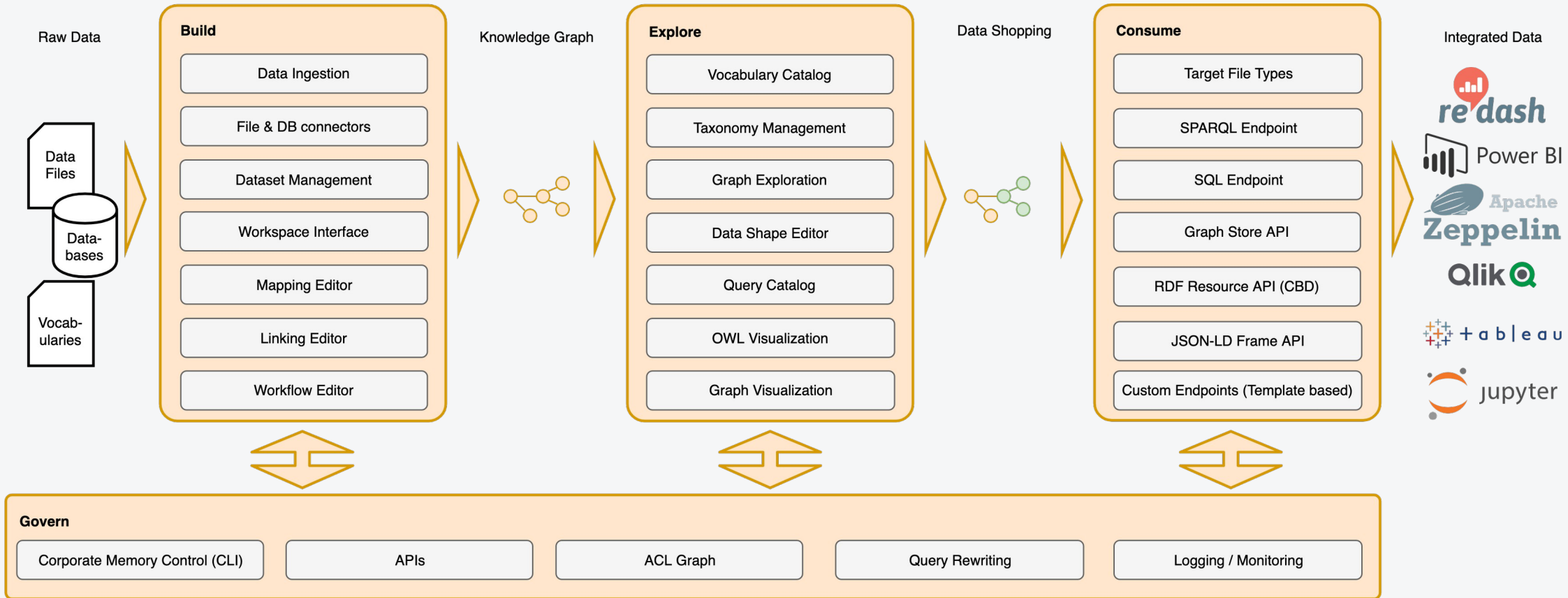
Given Name	Name	City	Profession	Organization
John	Doe	Berlin	Accountant	CoCoNut Cie
Eric	Molière	Paris	Writer	Hachette
Vinay	Kular	London	Producer	PilmCo
Bruce	Lee	Beijing	Actor	

First Name	Last Name	Address	Company
John	D	Unter den Linden, Berlin	CoCoNut Cie
Eric	Molière	Champs Elysées, Paris	
Vinay	Kumar	Richmond Hill, London	FilmCo
Brice	Lee		

Link?



User Journey and Functional Areas



Linking as general rules mechanism

eccenca Corporate Memory

link-loans-2-customer-1

Finished in 4.901s

Linkage Rule Fitness: 99.5 (based on 3 positive and 0 negative reference links)

Uncertain links

CONFIRM DECLINE NOT SURE DONE

Path	Value	Path	Value
rdfs:label	Aaron Yundt: 21937 USD	sdo:legalName	Aaron Yundt
sdo:birthDate	1974-07-10	sdo:birthDate	1974-07-10
rdf:type	http://schema.org/LoanOrCredit	rdf:type	http://schema.org/Rating
sdo:legalName	Aaron Yundt	sdo:postalCode	6564
sdo:interestRate	Variable	sdo:birthDate	1974-07-10
sdo:accountId	FUGBH86GF655	sdo:ratingValue	10
sdo:amount	21937	rdfs:label	Aaron Yundt: 10
sdo:currencyCode	USD		
sdo:postalCode	6564		

v20.03-6450-g2b14c25cc ©2020 eccenca GmbH

eccenca Corporate Memory

link-similar-products

Link Limit: unlimited Link Type: pv:similarProduct

No score available

Source Paths: Products-Demo-Int... (custom path) pv.height_mm Path (Source) pv.areaOfExpertise pv.id pv.hasProductManager iworice

Target Paths: Products-Demo-Int... (custom path) rdfs:label pv.id pv.hasProductManager iworice

Transformations recommended: Constant Lower case Tokenize

Comparators recommended: Jaccard Levenshtein distance Numeric Equality String Equality

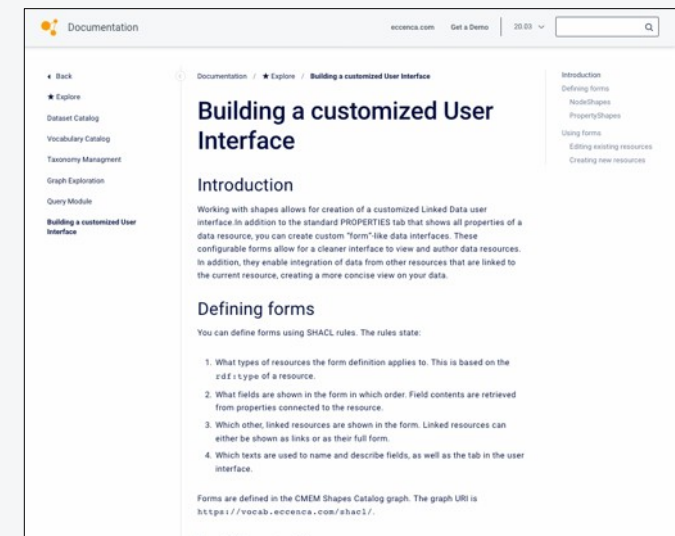
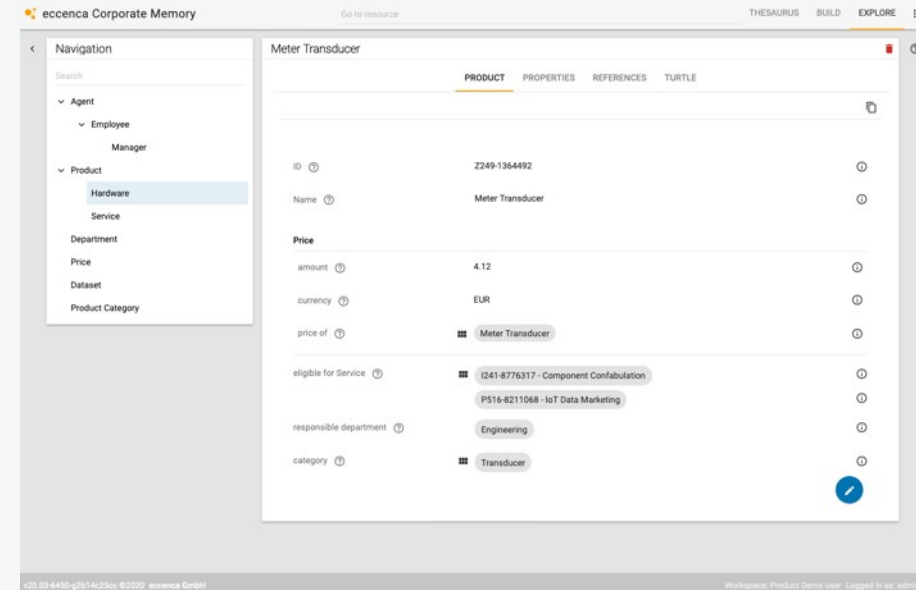
Aggregators recommended: And Average Or

v20.03-6450-g2b14c25cc ©2020 eccenca GmbH

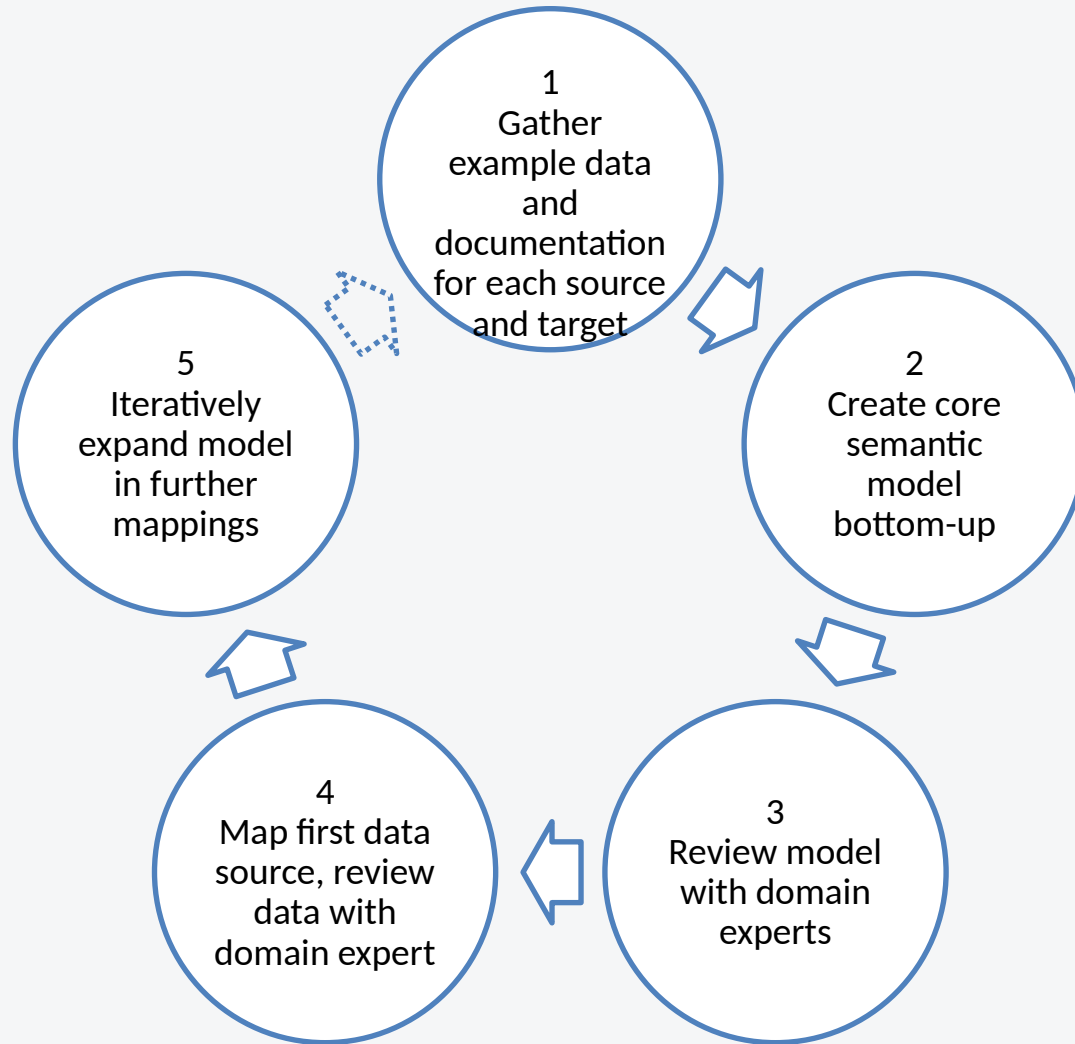
Data defined UI Configuration

- 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>

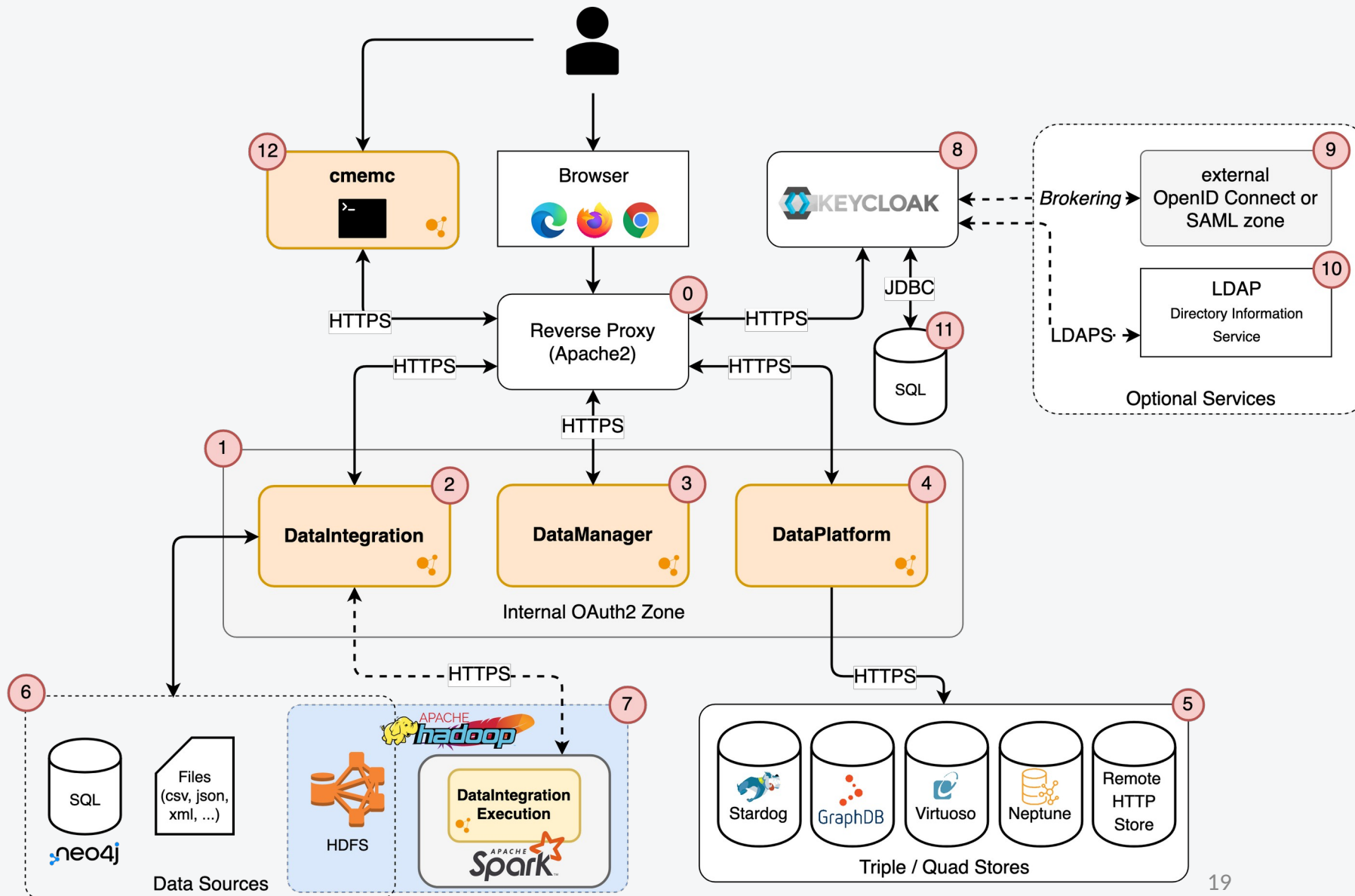


Enterprise Knowledge Graph Project Flow



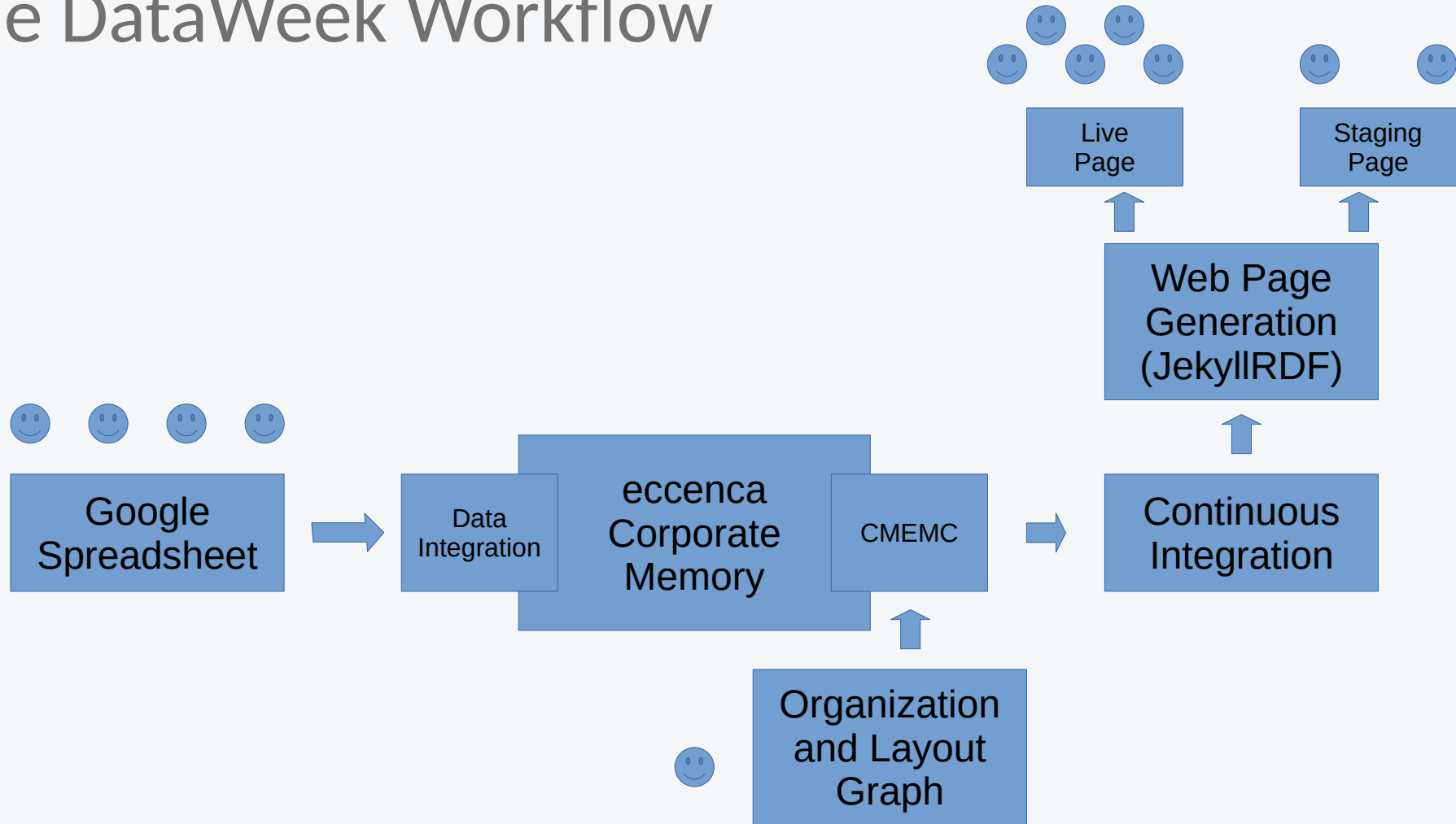
1. Analyze data sources and understand domain
 2. Find or build a semantic model (ontology/vocabulary)
 3. Create the mappings
 4. Clean, enrich and link data
 5. Explore the results
- ... Re-iterate if needed ...

Reference Architecture Corporate Memory



SAMPLE USE CASES

The DataWeek Workflow



The DataWeek Workflow

View Download View X

DataWeek Programm_Website Freigeben Anmelden

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

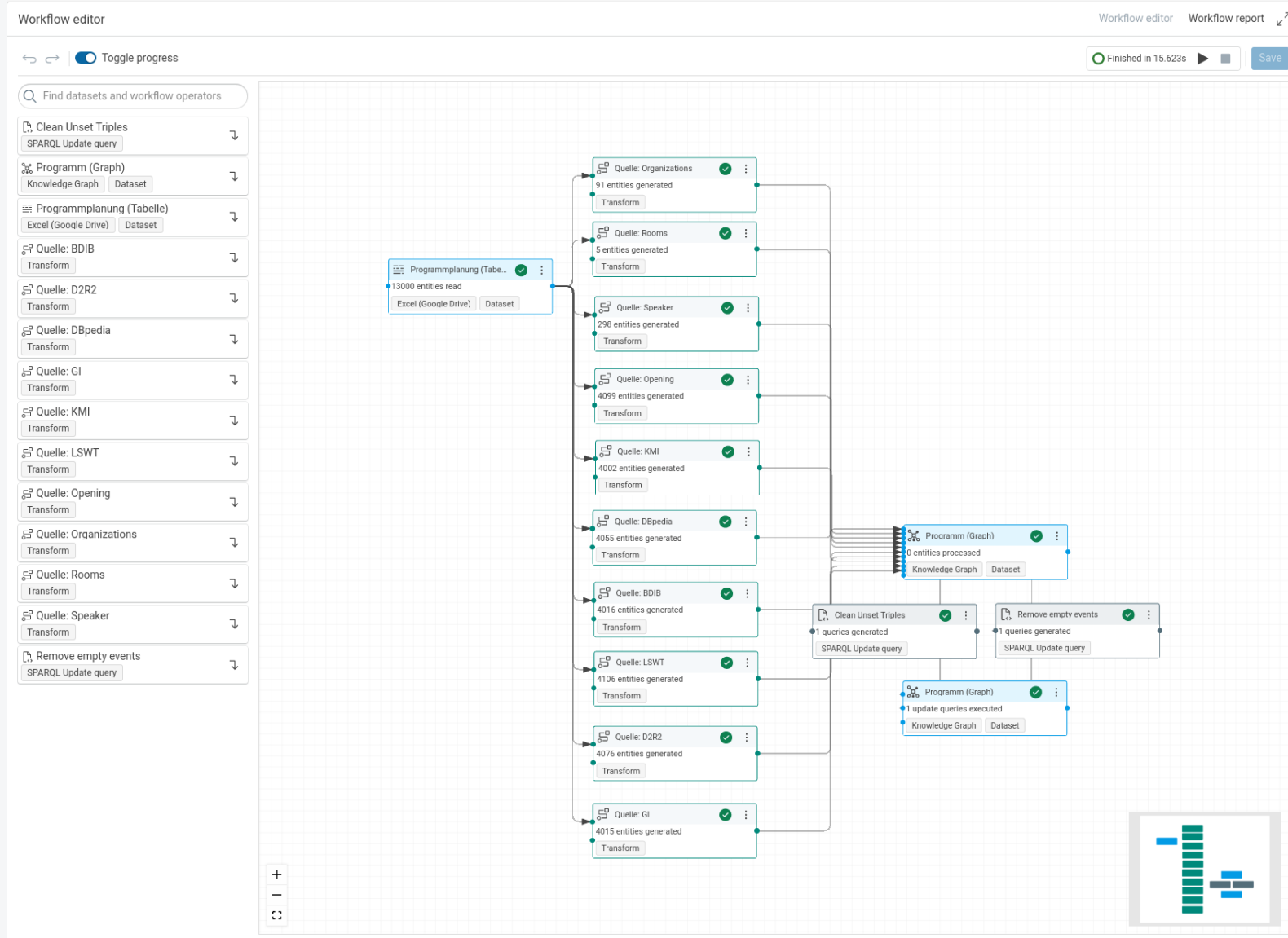
100% Nur Lesezugriff

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/#MehdiAzarafa		Mehdi Azarafa	Scientific Researcher	Wissenschaftlicher Mitarbeiter	https://infaif.org/
55	http://aksw.org/MichaelMartin	Dr.	Michael Martin	Head of Competence Center	Leiter Kompetenzzentrum	https://infaif.org/
56	http://aksw.org/MilanDojchinovski	Dr.	Milan Dojchinovski	Research Associate / Assistant Professor	Wissenschaftlicher Mitarbeiter / Assistenzprofessor	https://infaif.org/
57	https://dataweek.de/#MiriamWelz		Miriam Welz			https://scads.ai/
58	https://dataweek.de/#MirkoMuehlport		Mirko Mühlport	Team leader at City of Leipzig/Digital City Department, CUT Project	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://infaif.org/
60	https://dataweek.de/#NadineKuhlavonBergmann	Dr.	Nadine Kuhla von Bergmann	CEO	CEO	https://creativeclimaticities.org/de/
61	https://dataweek.de/#NadineWeissmann	Dr.	Nadine Weissmann	Requirements Engineer	Anforderungsingenieurin	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/digitale
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://infaif.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/#PauloRicardoViviurkadoCarvalho		Paulo Ricardo Viviurka do Carmo	Master student	Master student	https://infaif.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://infaif.org/
74	https://dataweek.de/#RuslanHrushchak		Ruslan Hrushchak	Managing director	Geschäftsführer	https://applant.de/de/home/
75	http://aksw.org/SabineGründerFahrer	Dr.	Sabine Gründer-Fahrer			https://infaif.org/
76	https://dataweek.de/#SaideshwarKota		Saideshwar Kota			https://www.wichita.edu
77	https://dataweek.de/#SakylMannah		Sakyl Mannah			https://srw.de/
78	https://dataweek.de/#SebastianGemkow		Sebastian Gemkow	Saxon State Minister for Science	Sächsischer Staatsminister für Wissenschaft	https://www.srnwk.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/digitale
80	http://aksw.org/SebastianHellmann	Dr.-Ing.	Sebastian Hellmann	AKSW, DBpedia Association	AKSW, DBpedia Association	https://infaif.org/
81	https://dataweek.de/#SebastianLemann	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://srw.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://infaif.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












Summary 

Label Quelle: Speaker

Mapping editor Mapping editor Transform evaluation Transform execution ↗

Person

- homepage
Document
- Affiliation
n/a

Person			⚙
 Person	Person, Person	{IRI}	▼
Mapping rules (9)			
 Description (de) <small>dcterms:description</small>	LanguageValueType (de)	Description_de	▼
 Description (en) <small>dcterms:description</small>	LanguageValueType (en)	Description_en	▼
 Homepage <small>foaf:homepage</small>	Document	n/a	→
 Pref Label <small>skos:prefLabel</small>	StringValue Type	Title, Name	▼
 Title <small>foaf:title</small>	StringValue Type	Title	▼
 Name <small>foaf:name</small>	StringValue Type	Name	▼
 Bio (de) <small>.../standards/elementset/gnd#biographicalOrHistoricalInformation</small>	LanguageValueType (de)	Bio_de	▼
 Bio (en) <small>.../standards/elementset/gnd#biographicalOrHistoricalInformation</small>	LanguageValueType (en)	Bio_en	▼
 Affiliation <small>schema:affiliation</small>	n/a	Affiliation	→ 

The DataWeek Workflow

The screenshot displays the DataWeek workflow interface. At the top, the user profile for Dr. Sebastian-Philipp Brandt is shown. The main content area is divided into several sections:

- Navigation:** A sidebar on the left contains a search bar and a list of categories: Person, Document, Organization, di_Dataset, Event, and dataweek.de_Keynote.
- Resource Properties:** The main area shows the properties of the resource "Dr. Sebastian-Philipp Brandt". The properties are listed in a table with columns for the property name, the value, and edit/delete icons.

Property	Value	Actions
type	Person	[edit] [delete]
	Person	[edit] [delete]
	SHOW IN LIST ADD	
Description	Senior Key Expert: Knowledge Graph and Data Management@de	[edit] [delete]
	Senior Key Expert: Knowledge Graph and Data Management@en	[edit] [delete]
	SHOW IN LIST ADD	
preferred label	Dr. Sebastian-Philipp Brandt	[edit] [delete]
	SHOW IN LIST ADD	
affiliation	Siemens AG	[edit] [delete]
	SHOW IN LIST ADD	
name	Sebastian-Philipp Brandt	[edit] [delete]
	SHOW IN LIST ADD	
title	Dr.	[edit] [delete]
	SHOW IN LIST ADD	
gnd_biographicalOrHistoricalInformation	Sebastian Brandt is Senior Key Expert for knowledge-graphs and data management at Siemens Technology, the global research and development centre of Siemens. Apart from research, Sebastian is helping Siemens businesses get most out of their data across the entire product life-cycle. This involves the creation and maintenance of industrial knowledge-graphs and their deployment in real-world production scenarios. Prior to Siemens, Sebastian has worked extensively on ontology-based knowledge representation systems for the life-sciences in the Bio-Health Informatics Group at the University of Manchester. His research on subsets of OWL with efficient inference problems has opened the door to the OWL-EL facet in the ontology standard OWL 2.@en	[edit] [delete]
	Sebastian Brandt ist Senior Key Expert für Wissensgraphen und Datenmanagement bei Siemens Technology, dem globalen Forschungs- und Entwicklungszentrum von Siemens. Neben seiner Tätigkeit in der Forschung unterstützt Sebastian Brandt Siemens-Unternehmen dabei, das Beste aus ihren Daten über den gesamten Produktlebenszyklus hinweg herauszuholen. Dies beinhaltet die Erstellung und Pflege von industriellen Wissensgraphen und deren Einsatz in realen Produktionsszenarien. Vor seiner Tätigkeit bei Siemens hat Sebastian in der Bio-Health Informatics Group an der University of Manchester intensiv an ontologiebasierten Wissensrepräsentationssystemen für die Biowissenschaften gearbeitet. Seine Forschung zu Teilmengen von OWL mit effizienten Inferenzproblemen hat die Tür zur OWL-EL-Facetten im Ontologie-Standard OWL 2 geöffnet.@de	[edit] [delete]
	SHOW IN LIST ADD	

Found 7 results Properties per page 20 |<< Page 1 of 1 >>



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

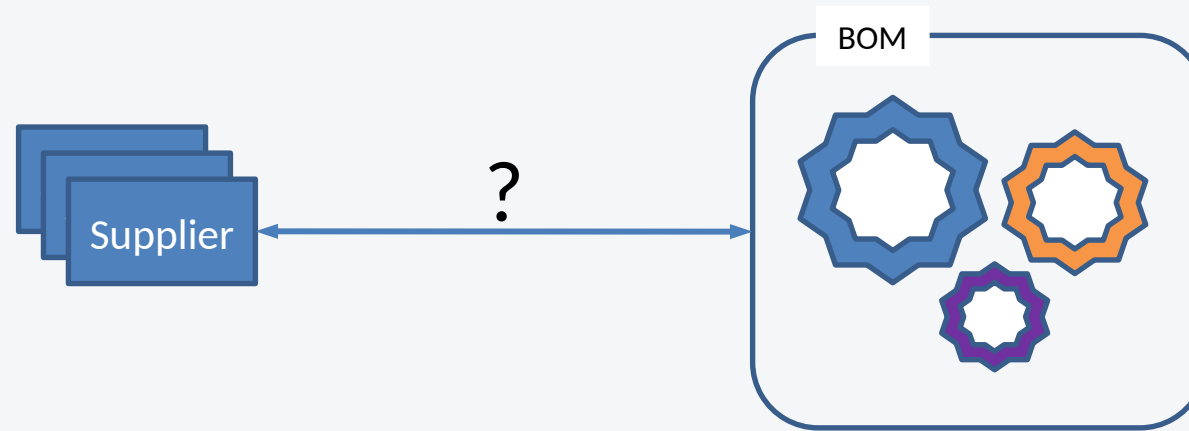
Hainstraße 8
D-04109 Leipzig
Germany

+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:

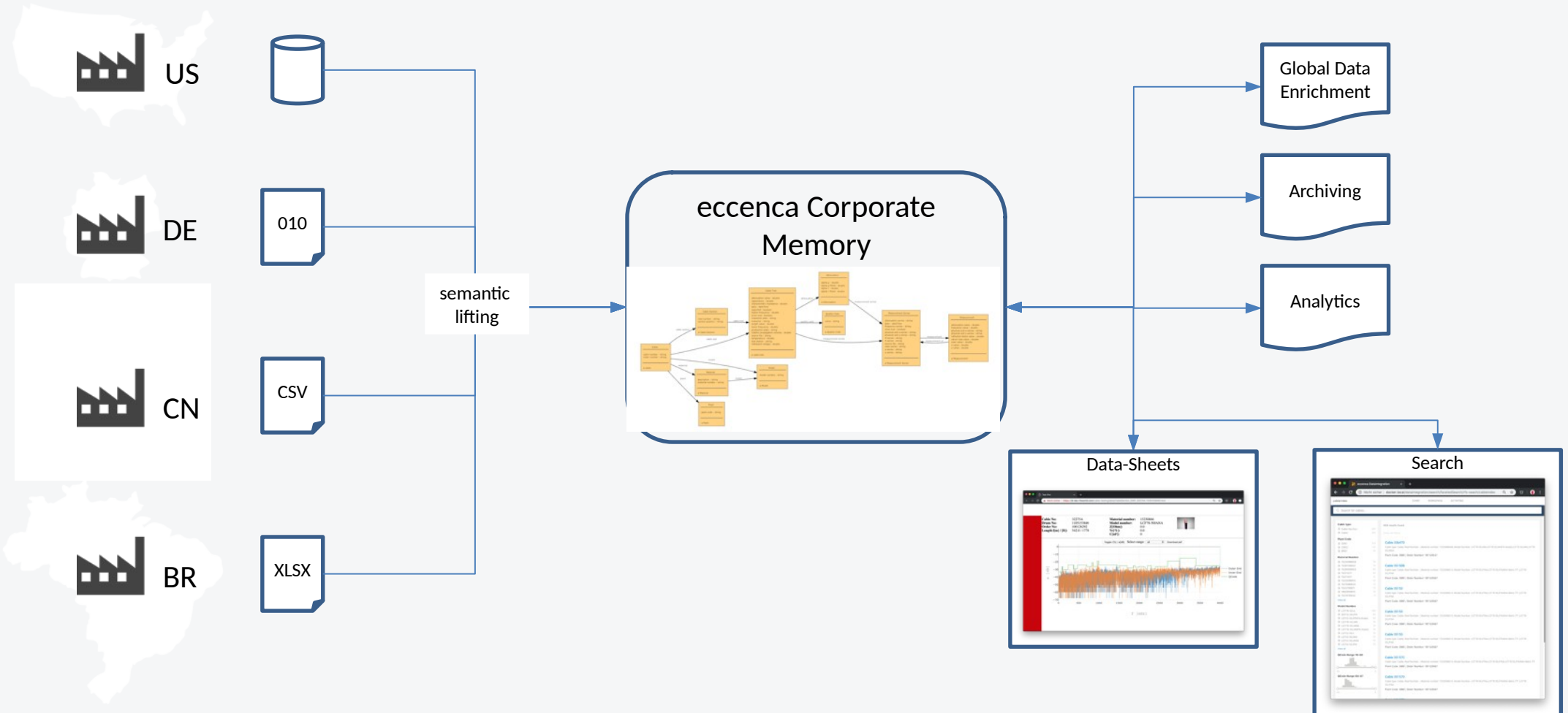
MPN	Supplier
293D105X9016A2##E3	VISHAY
B45196E3105K10	KEMET
T491A105K016AS	KEMET
T491A105K016AT	KEMET
TAJA105K016RNJ	AVX

Machine learning based on granular material properties like:

Capacitor Type	TANTALUM CAPACITOR
Capacitance	1.0 μ F
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



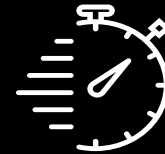
Project Benefits

- 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

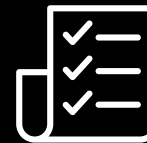
BENEFIT Cases



- Global S&OP
Capacity Planning
- Less global inventory
 - Better factory balancing



- Customer Service Teams
- Improvement lead-times
 - Automation



- Customer Experience
- Self-service, on-demand production data

About eccenca GmbH

brox IT-Solutions GmbH



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



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

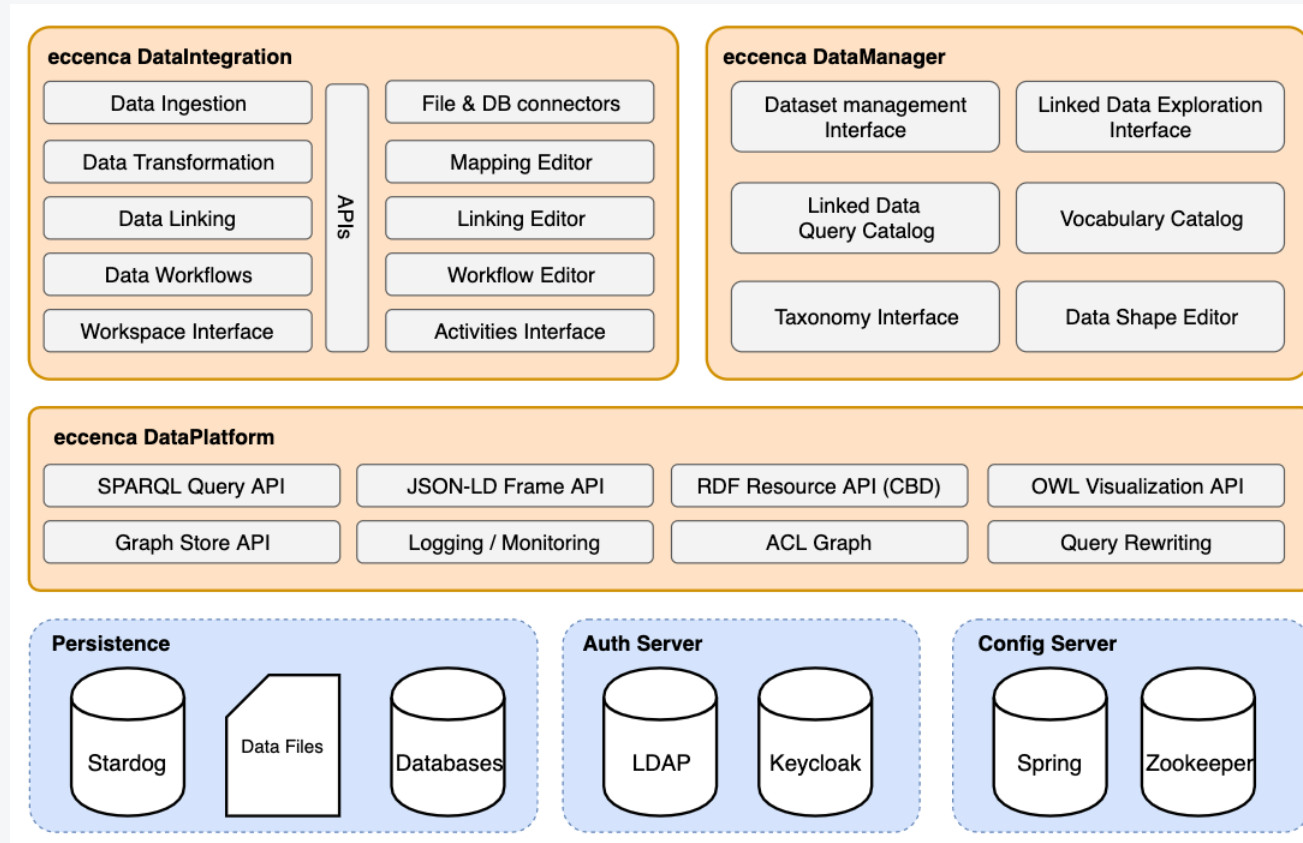
Initiator: DBPedia, Linked GeoData etc.
Betreiber: Datenportal der EU-Kommission
Team-Leitung: Prof. Dr. Sören Auer

eccenca GmbH



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

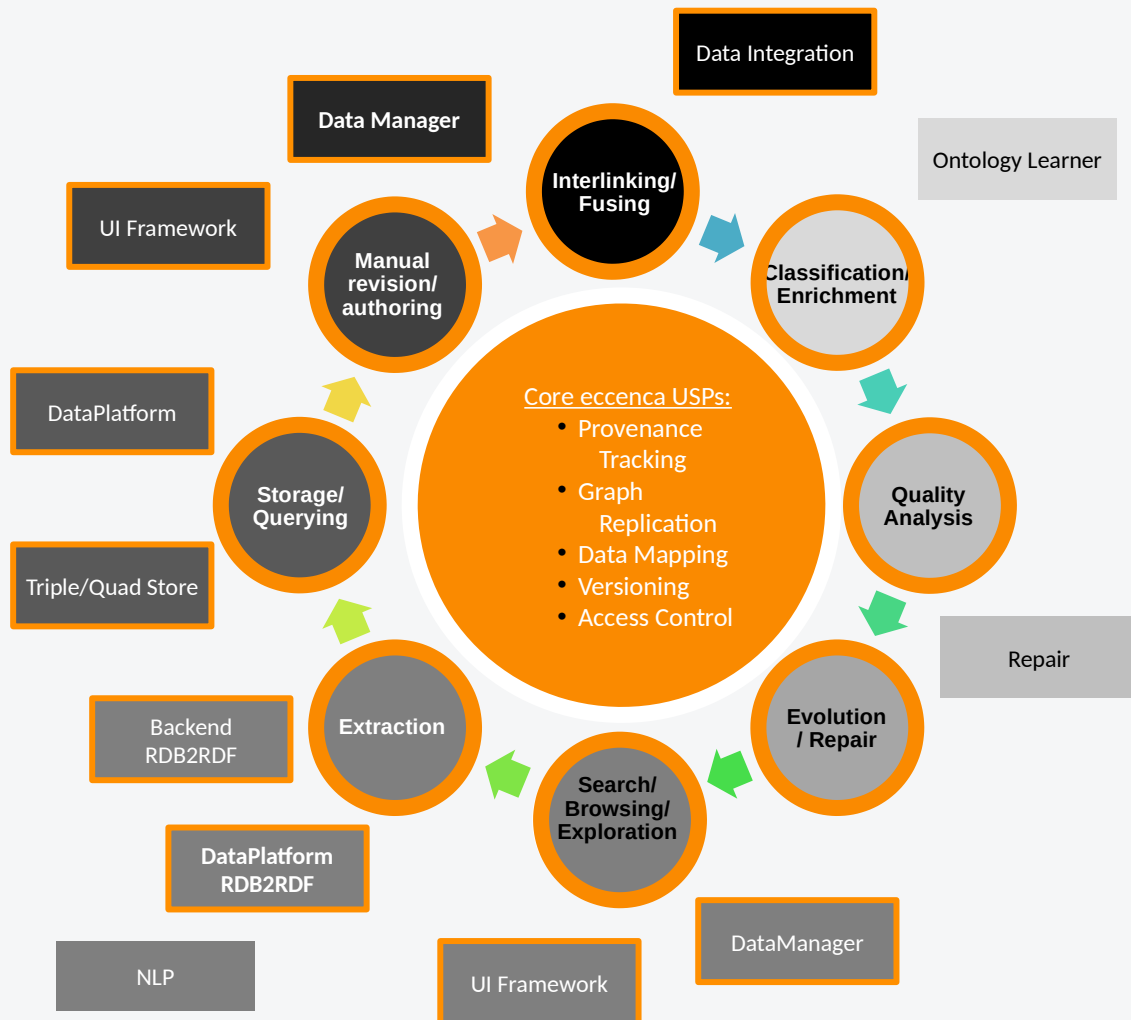
What

- § 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

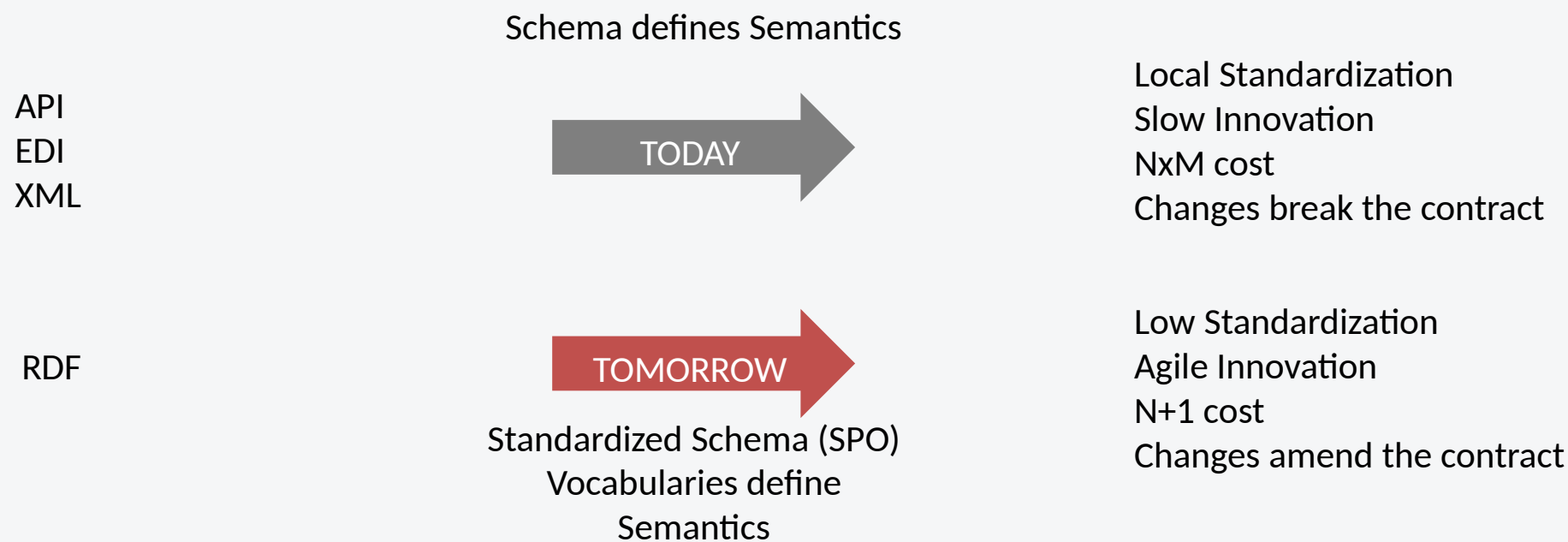
Linked Data Life Cycle



- 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



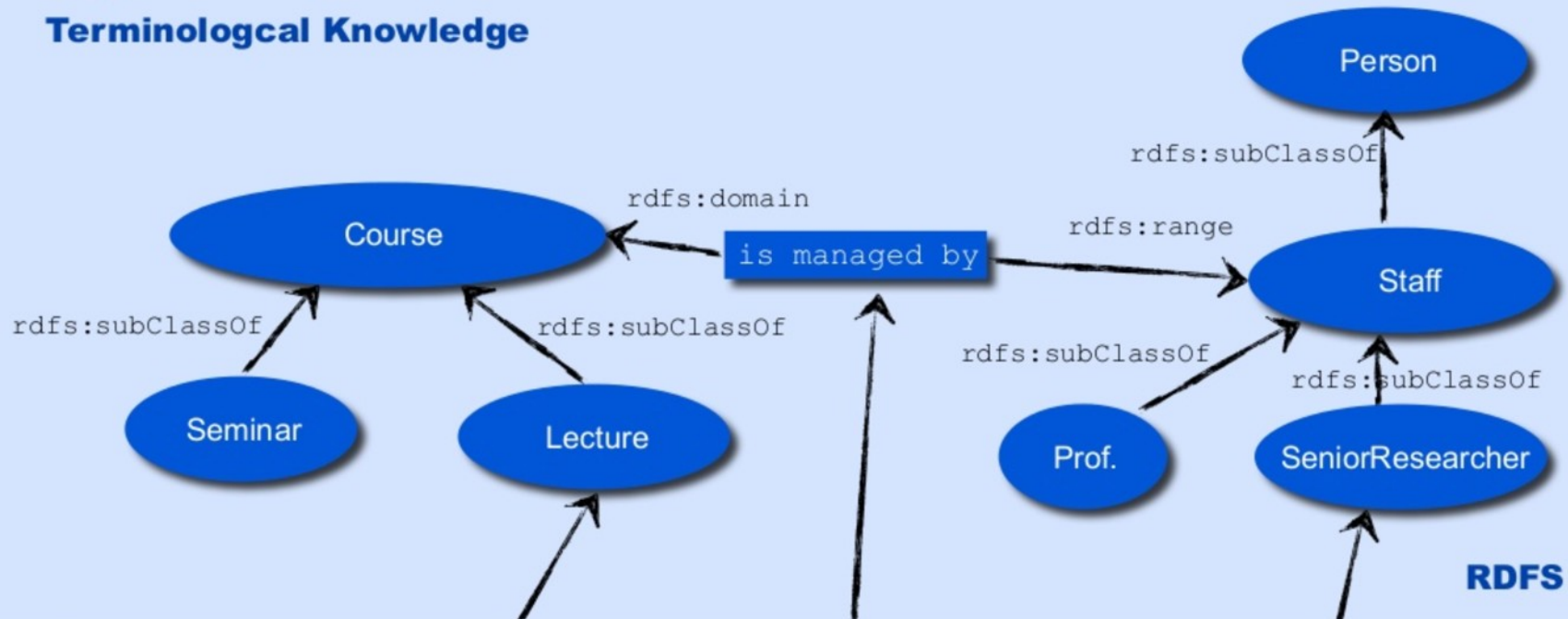
Other Use Cases

- 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 Monetization Platform project
- Semantic Enterprise Information Model
 - MDM, simplification, data preparation, analytics, etc.

RDF + RDFS

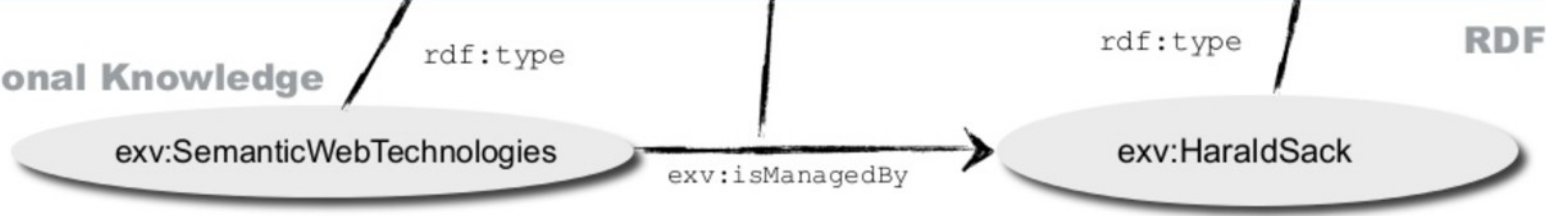
13

T-Box Terminological Knowledge



RDFS

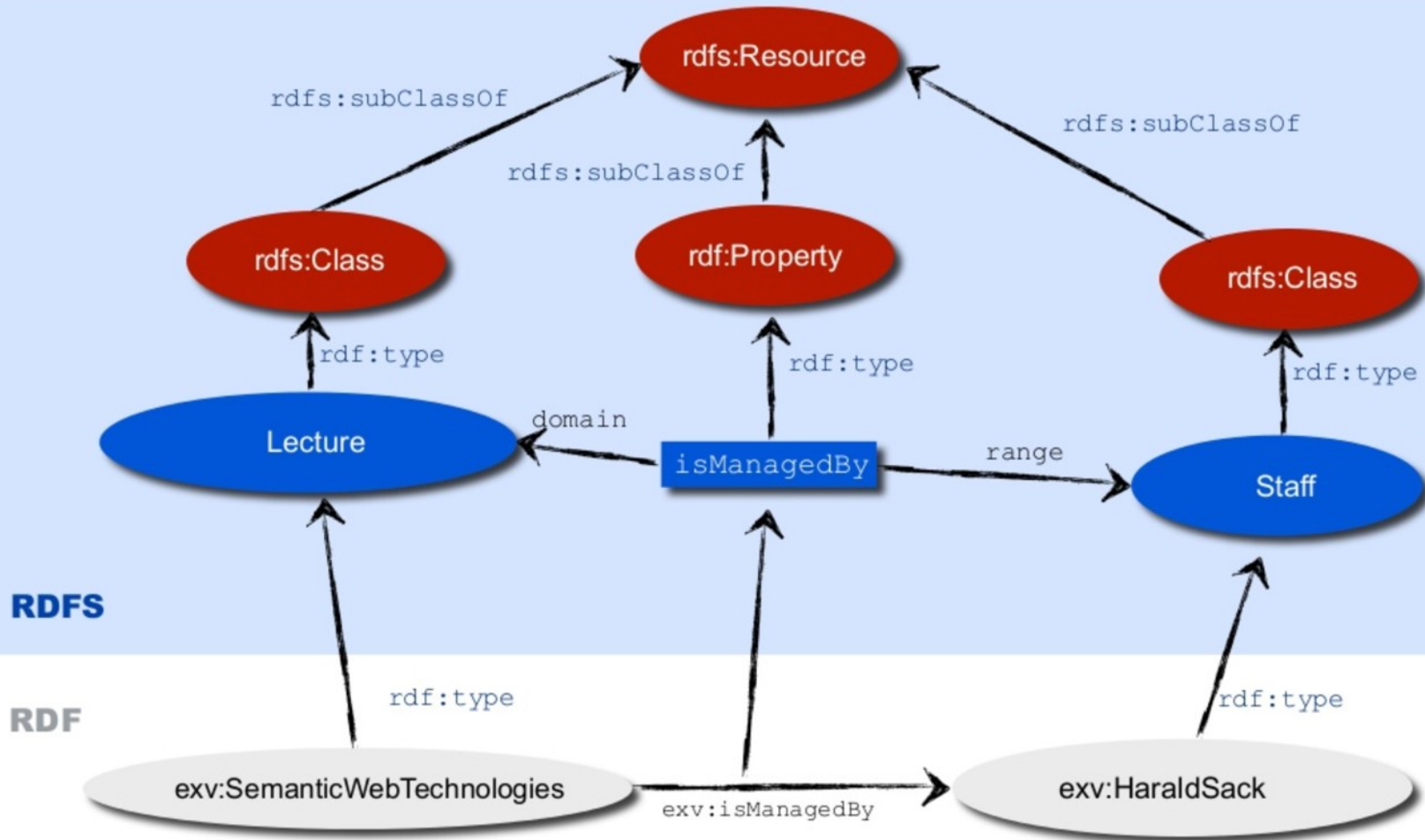
A-Box Assertional Knowledge



RDF

Classes, Properties, Instances

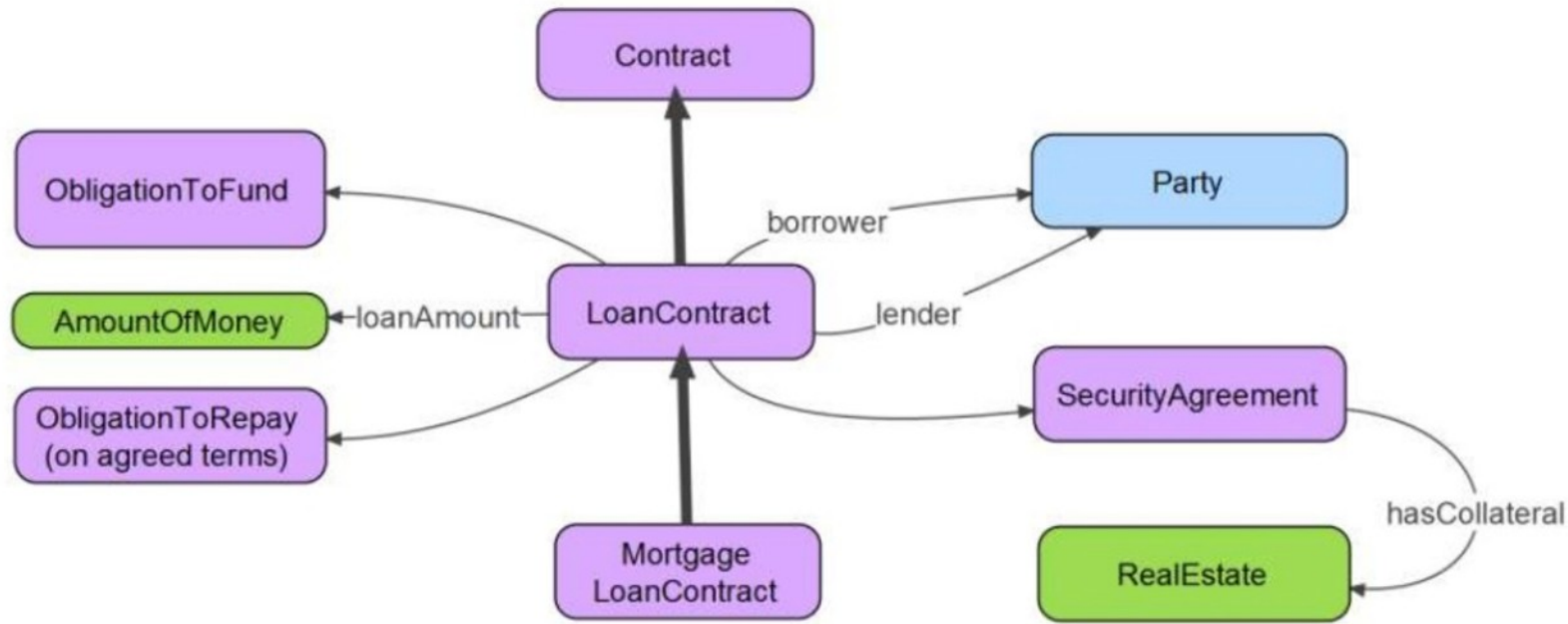
8



RDFS

RDF

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 .
```

