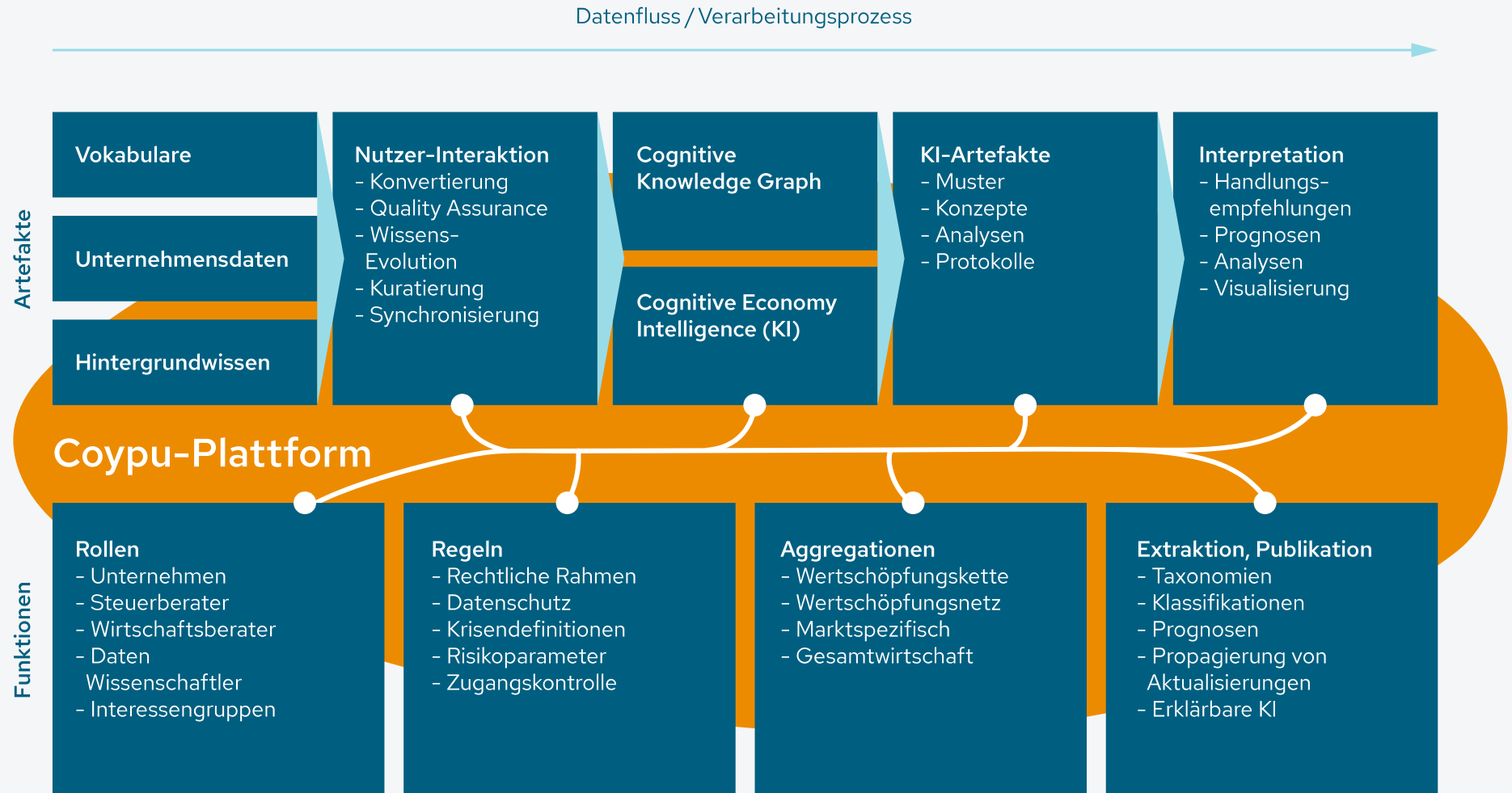




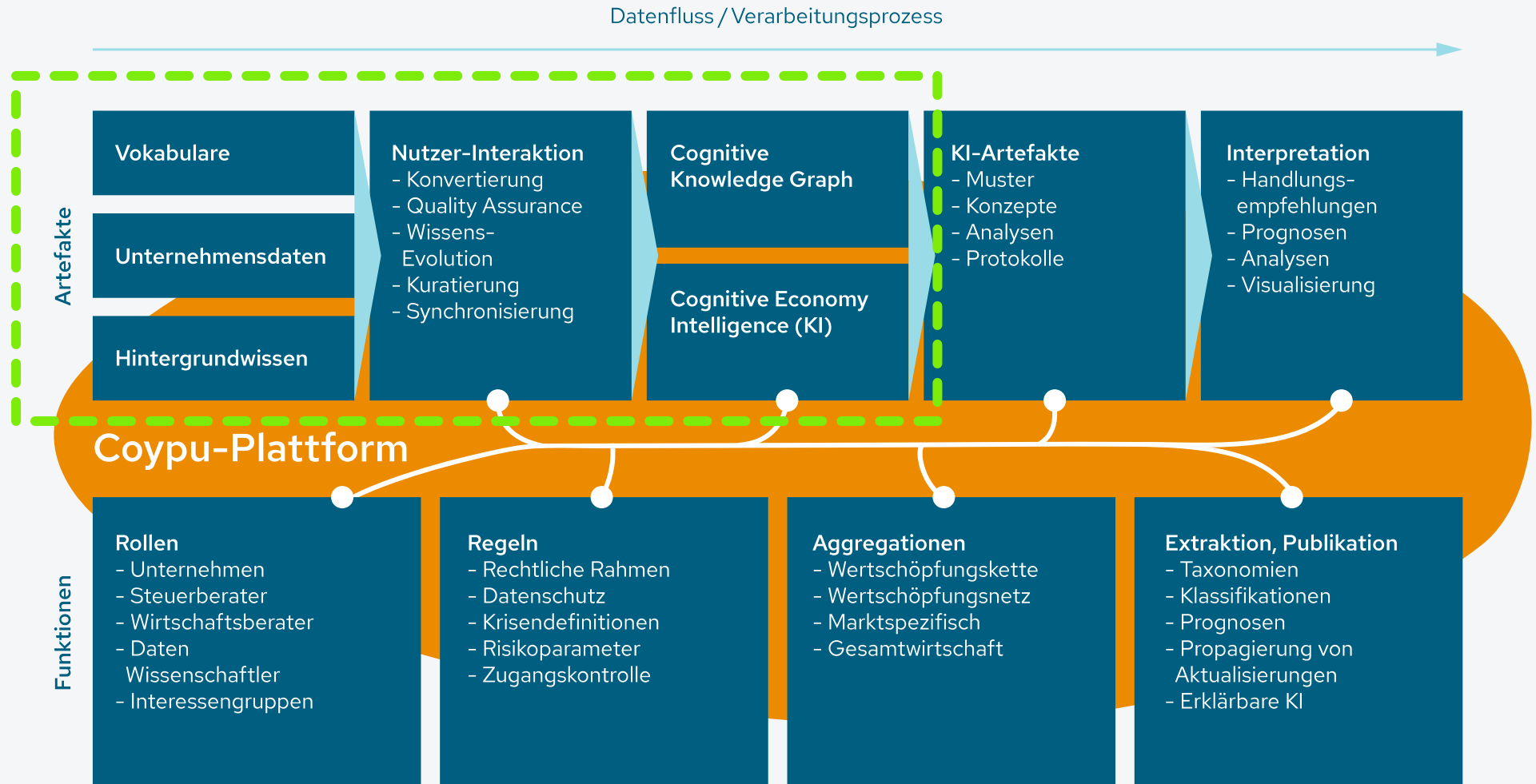
The Coypu Platform Data Integration Workflow

Dr. Natanael Arndt
Senior Linked Data Expert @ eccenca GmbH

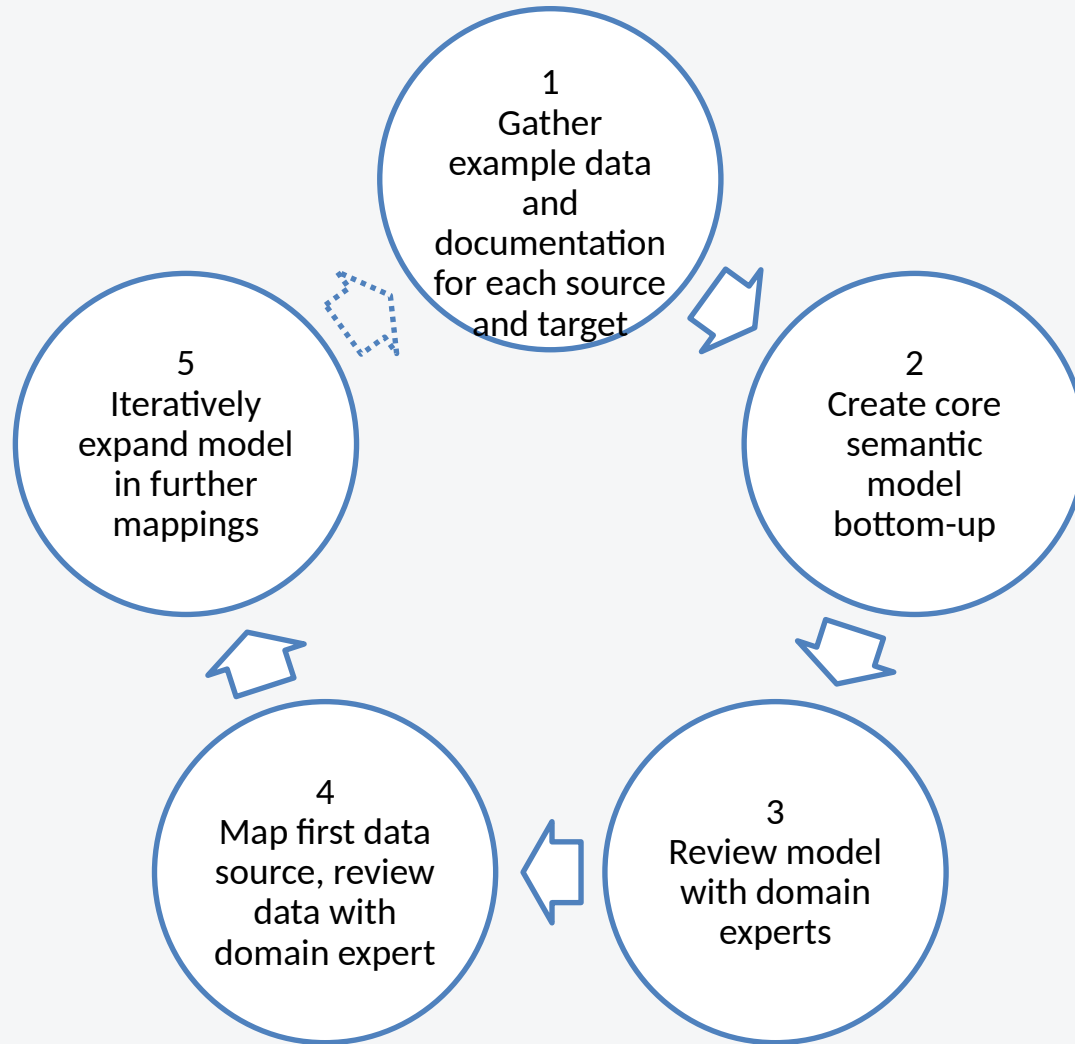
The Coypu-Plattform



The Coypu-Plattform



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 ...

Data Sources

Data.europa.eu English (en) Search datasets

Data Studies data.europa academy News Contact

Datasets Feed

Datasets

Filter by location Search datasets

Leaflet | © OpenStreetMap contributors © GISCO

Datasets Catalogues Editorial Content Last Modified Relevance More

1 425 122 datasets found

Oficiālā elektroniskā adrese (e-adrese)

Informācija par Oficiālās elektroniskās adreses aktivizētajiem, deaktivizētajiem kontiem un anulētajām e-adresēm. Dati tiek atjaunoti reizi stundā. CSV

Updated: 04.07.2022 18:30 Created: 01.10.2019 11:19 data.gov.lv

COVID-19 apstiprināto gadījumu skaits un 14 dienu kumulatīvā saslimstība pa administratīvajām teritorijām

Slimību profilakses un kontroles centram iesniegtā operatīvā informācija par apstiprinātajiem COVID-19 gadījumiem. Dati tiek apkopoti reizi dienā no laboratorijām un ārstniecības personām. Iepriekš publicētie dati sadalījumā pa administratīvajām teritorijām var tikt aktualizēti atbilstoši precīzētai informācijai par personas dzīvesvietu. No 21.05.2020. līdz... CSV JSON

Updated: 04.07.2022 16:22 Created: 08.04.2020 02:00 data.gov.lv

COVID-19 izmeklējumi, apstiprinātie gadījumi un iznākumi

GeoNames

The GeoNames geographical database covers all countries and contains over eleven million placenames that are available for download free of charge.

Search all countries

search advanced search

enter a location name, ex: "Paris", "Mount Everest", "New York"

Browse the names

- Countries
- Postal codes
- Country statistics
- Recent modifications

Information

- About GeoNames
- Data Sources
- User manual
- Ambassadors and Team
- Forum
- Blog
- Mailing list
- Commercial Support and Consulting

Download

- Info
- Free Gazetteer Data
- Free Postal Code Data
- Premium Data

Web Services

- Overview
- Documentation
- Client Libraries
- Premium Web Services

Sponsoring

- Nettikasino
- Nye Casino Online
- Neue Online Casinos
- Japanesecasino.com
- Casinos not on Gamstop
- nuovicasino.it
- BestUK.Casino
- Mäklare
- SveaCasino.se
- CasinoFeberse
- Casino utan Spelpaus
- Pelisivut.com
- Suomicasino.com
- online casino
- Casino Frog
- Ruletti
- Donations and Sponsoring
- uudet kasinot
- Nettikasino.org
- Casino utan Svensk Licens
- casino på nett
- Michigan Sports Betting
- casinonlinechile.com
- UK casinos not on GamStop
- Alcasino.se

This work is licensed under a [Creative Commons Attribution 4.0 License](#)

Authoritative Country Dataset from the EU
XML Dataset

Open CC-BY licensed global geo information
CSV/TSV Files

The Vocabulary

The image displays two screenshots of a web interface for the 'country' class in the CoyPu Ontology. The left screenshot shows the 'Properties' tab, and the right screenshot shows the 'References' tab.

Left Screenshot (Properties):

- Resource: **country**
- Navigation: **Class** (selected), AnnotationProperty, Datatype, ObjectProperty, DatatypeProperty, FunctionalProperty, > crm entity
- Properties:
 - type**: Class (SHOW IN LIST, ADD)
 - subClassOf**: location (SHOW IN LIST, ADD)
 - Label**: country^{en} (SHOW IN LIST, ADD)
 - Comment**: a country as defined in ISO 3166^{en} (SHOW IN LIST, ADD)
 - seeAlso**: ontology_Country, schema.org_Country (SHOW IN LIST, ADD)
- Found 5 results Properties per page 20 | < < Page 1 of 1 > >

Right Screenshot (References):

- Resource: **country**
- Navigation: **Class** (selected), AnnotationProperty, Datatype, Ontology, ObjectProperty, DatatypeProperty, FunctionalProperty, > crm entity
- References Table:

Subject t_1	Predicate t_1	Graph t_1
has adaptive capability	domain	coy: CoyPu Ontology
has coping capability	domain	coy: CoyPu Ontology
has currency	domain	coy: CoyPu Ontology
has exposure	domain	coy: CoyPu Ontology
has ISO code	domain	coy: CoyPu Ontology
has language	domain	coy: CoyPu Ontology
has population	domain	coy: CoyPu Ontology
has risk level	domain	coy: CoyPu Ontology
has susceptibility	domain	coy: CoyPu Ontology
has vulnerability	domain	coy: CoyPu Ontology
- Found 10 results. Rows per page 20 | < < Page 1 of 1 > >

TODO: publish under <https://schema.coypu.org/global>

Transformation of the Datasources

Summary ✎

Label publications.europa.eu Extract Country XML distribution

Description TRANSFORMATION 1 in <https://ns.coypu.org/country> ... more

Created by [unknown user](#). Last modified 2022/06/07 by [white-gecko](#).

Mapping editor Mapping editor Transform evaluation Transform execution ↗

- ▼ root
- URI of code-3166-1-alpha- n/a
- URI of code-iana n/a
- URI of code-3166-1-num n/a
- URI of authority-code n/a
- URI of code-3166-1-alpha-1 n/a
- URI of code-3166- n/a
- URI of UNSD-geoscheme n/a
- URI of code- M n/a
- EU country concept n/a

root ☰

root Country <https://data.coypu.org/country/{code-3166-1-alpha-3}> ^

Multiple entities are allowed

Target entity type
Country · coy:Country · a country as defined in ISO 3166

URI pattern
<https://data.coypu.org/country/{code-3166-1-alpha-3}> ✎ 🗑️

Examples of target data

Value path	Value	Transformed value
/code-3166-1-alpha-3	AND	https://data.coypu.org/country/AND

Label
root

Edit Copy

Mapping rules (19)

code-3166-1-alpha- coy.code-3166-1-alpha-2	StringValue Type	code-3166-1-alpha-2	▼
URI of code-3166-1-alpha- owl.sameAs	n/a	n/a	→
code-iana coypu.ref.code-iana	StringValue Type	code-iana	▼
URI of code-iana owl.sameAs	n/a	n/a	→
code-3166-1-num coy.code-3166-1-num	StringValue Type	code-3166-1-num	▼
URI of code-3166-1-num owl.sameAs	n/a	n/a	→
authority-code coy.authority-code	StringValue Type	authority-code	▼
URI of authority-code owl.sameAs	n/a	n/a	→
code-3166-1-alpha-1	StringValue Type	code-3166-1-alpha-3	▼

Transformation of the Datasources

Summary

Label: publications.europa.eu Extract Country XML distribution

Description: TRANSFORMATION 1 in <https://ns.coypu.org/country> ... more

Created by unknown user. Last modified 2022/06/07 by white-gecko.

Mapping editor

- root
 - URI of code-3166-1-alpha- n/a
 - URI of code-iana n/a
 - URI of code-3166-1-num n/a
 - URI of authority-code n/a
 - URI of code-3166-1-alpha-1 n/a
 - URI of code-3166- n/a
 - URI of UNSD-geoscheme n/a
 - URI of code- M n/a
 - EU country concept n/a

root

Country <https://data.coypu.org/country/{code-3166-1-alpha-3}>

Multiple entities are allowed

Target entity type: Country · coy:Country · a country as defined in ISO 3166

URI pattern: <https://data.coypu.org/country/{code-3166-1-alpha-3}>

Examples of target data

Value path	Value	Transformed value
/code-3166-1-alpha-3	AND	https://data.coypu.org/country/AND

Label: root

Edit Copy

Mapping rules (19)

code-3166-1-alpha-coy:code-3166-1-alpha-2	StringValue Type	code-3166-1-alpha-2
URI of code-3166-1-alpha-owl:sameAs	n/a	n/a
code-iana-coy:ref:code-iana	StringValue Type	code-iana
URI of code-iana-owl:sameAs	n/a	n/a
code-3166-1-num-coy:code-3166-1-num	StringValue Type	code-3166-1-num
URI of code-3166-1-num-owl:sameAs	n/a	n/a
authority-code-coy:authority-code	StringValue Type	authority-code
URI of authority-code-owl:sameAs	n/a	n/a
code-3166-1-alpha-1	StringValue Type	code-3166-1-alpha-3

Mapping editor

Mapping editor Transform evaluation Transform execution

Label URI

Search term...

Source paths: Country XML distri...

- (custom path)
- @deprecated
- @id
- @IMMC.proposal.date
- @IMMC.approval.date
- @date.creation
- @adm.status
- @celex
- @pub
- code-3166-1-alpha-2

Transformations Recommended

- Constant
- Default Value
- Lower case
- Tokenize

Explore the Data

CoyPu Country and Location / Country list / Germany

Germany

Go to resource

User (18) Vocabularies (6) System (1)

https://data.coypu.org/country



























CoyPu Country and Location
https://data.coypu.org/country


Navigation

Search

- Country
- di_Dataset
- coy:TerritorialEntityOfSingleCountry
- coy:Country-EU

Properties Turtle

rdf:type	Country	SHOW IN LIST ADD	 
owl:sameAs	geoscheme_Western_Europe		 
	2_DE		 
	iana_de		 
	num_276		 
	code_DEU		 
		SHOW ALL 8 IN LIST ADD	
label	Germany	SHOW IN LIST ADD	 
coy:ref:code-iana	.de	SHOW IN LIST ADD	 
coy:code-3166-1-alpha-2	DE	SHOW IN LIST ADD	 
coy:code-3166-1-num	276	SHOW IN LIST ADD	 
coy:authority-code	DEU	SHOW IN LIST ADD	 
coy:code-3166-1-alpha-3	DEU	SHOW IN LIST ADD	 
coy:UNSD-geoscheme	Western Europe	SHOW IN LIST ADD	 



Explore the Data

CoyPu Country and Location / Country list / Germany
Germany

User (18)
Vocabularies (6)
System (1)

CoyPu Country and Location
https://data.coypu.org/country

Properties
Turtle

rdf:type	Country
	SHOW IN LIST ADD
owl:sameAs	geoscheme_Western_Europe
	SHOW IN LIST ADD
	2_DE
	SHOW IN LIST ADD
	iana_de
	SHOW IN LIST ADD
	num_276
	SHOW IN LIST ADD
	code_DEU
	SHOW ALL 8 IN LIST ADD
label	Germany
	SHOW IN LIST ADD
coypu-ref:code-iana	.de
	SHOW IN LIST ADD
coy:code-3166-1-alpha-2	DE
	SHOW IN LIST ADD
coy:code-3166-1-num	276
	SHOW IN LIST ADD
coy:authority-code	DEU
	SHOW IN LIST ADD
coy:code-3166-1-alpha-3	DEU
	SHOW IN LIST ADD
coy:UNSD-geoscheme	Western Europe
	SHOW IN LIST ADD

User (18)
Vocabularies (6)
System (1)

CoyPu Country and Location
https://data.coypu.org/country

Properties
Turtle

```

1 @prefix coy: <https://schema.coypu.org/global#> .
2 @prefix coypu-ref: <https://ns.coypu.org/reference/> .
3 @prefix owl: <http://www.w3.org/2002/07/owl#> .
4 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
5 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
6 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7
8 <https://data.coypu.org/country/DEU>
9   rdf:type          coy:Country ;
10  rdfs:label        "Germany" ;
11  owl:sameAs      <https://data.coypu.org/country/code-3166-1-num/276> , <https://data.coypu.
12  coypu-ref:code-iana ".de" ;
13  coy:UNSD-geoscheme "Western Europe" ;
14  coy:authority-code "DEU" ;
15  coy:code-3166-1-alpha-2 "DE" ;
            
```

[UPDATE](#)

User (18)
Vocabularies (6)
System (1)

CoyPu Country and Location
https://data.coypu.org/country

Navigation

Country

[di_Dataset](#)

[coy:TerritorialEntityOfSingleCountry](#)

[coy:Country-EU](#)

Linking

Linking editor

Linking editor | Linking evaluation | Linking execution | Reference links | Learning ↗

↶ ↷ | 🧩 | Show evaluation

🟢 Caches: 4 hours / 4 hours | ▶ | Save

All | 📄 | 🧩 | 🔍 | 📏 | ✨

🔍 Search for operators and pat

Source path ⓘ
Input | Recommended

Target path ⓘ
Input | Recommended

Lower case ⓘ
Transform | Normalize | Recomm

Tokenize ⓘ
Transform | Tokenization | Recom

Constant ⓘ
Transform | Value | Recommende

Default Value ⓘ
Transform | Value | Recommende

Levenshtein distance ⓘ
Comparison | Characterbased | Re

String equality ⓘ
Comparison | Equality | Recomm

Jaccard ⓘ
Comparison | Tokenbased | Recor

Numeric equality ⓘ

+

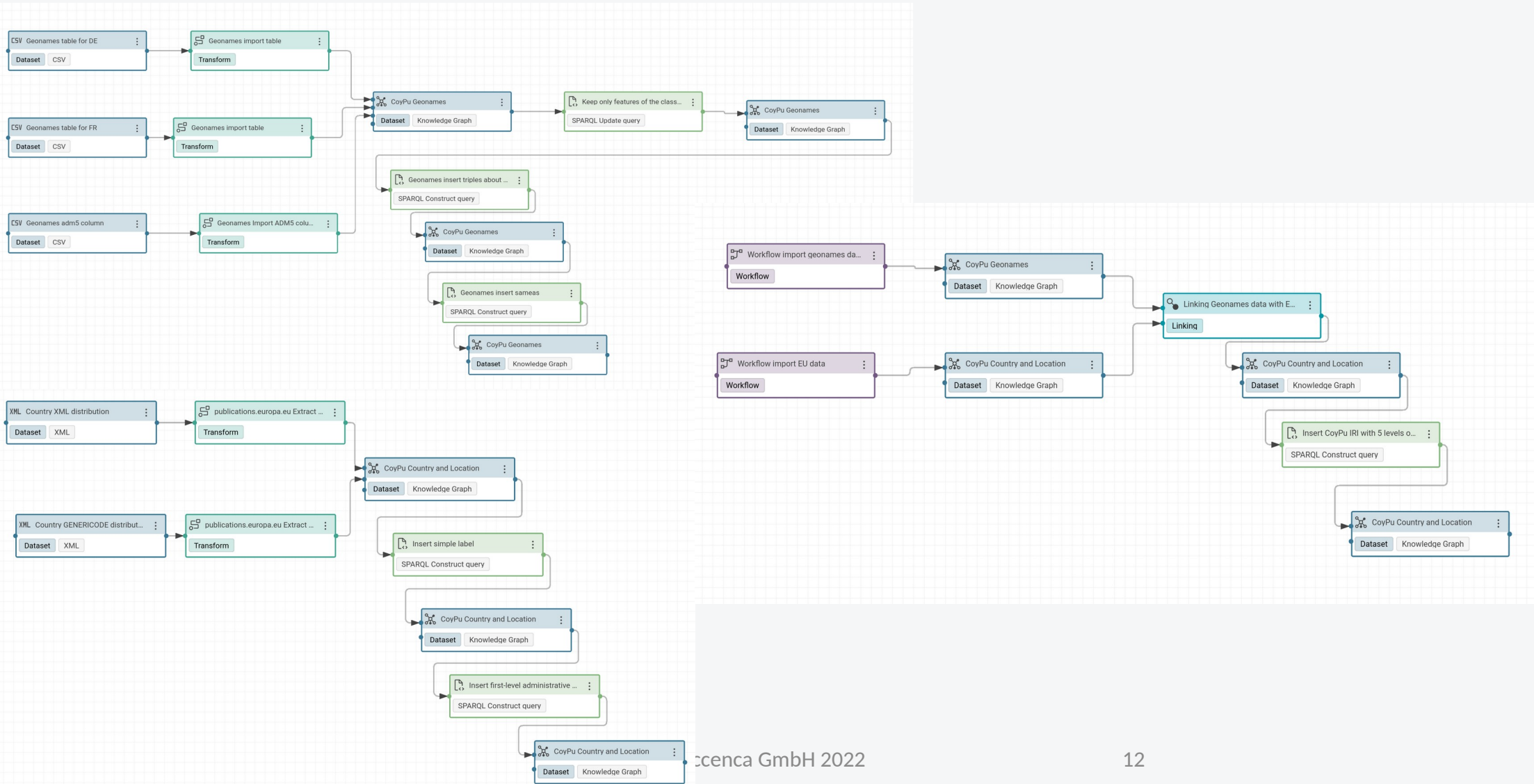
-

🔄

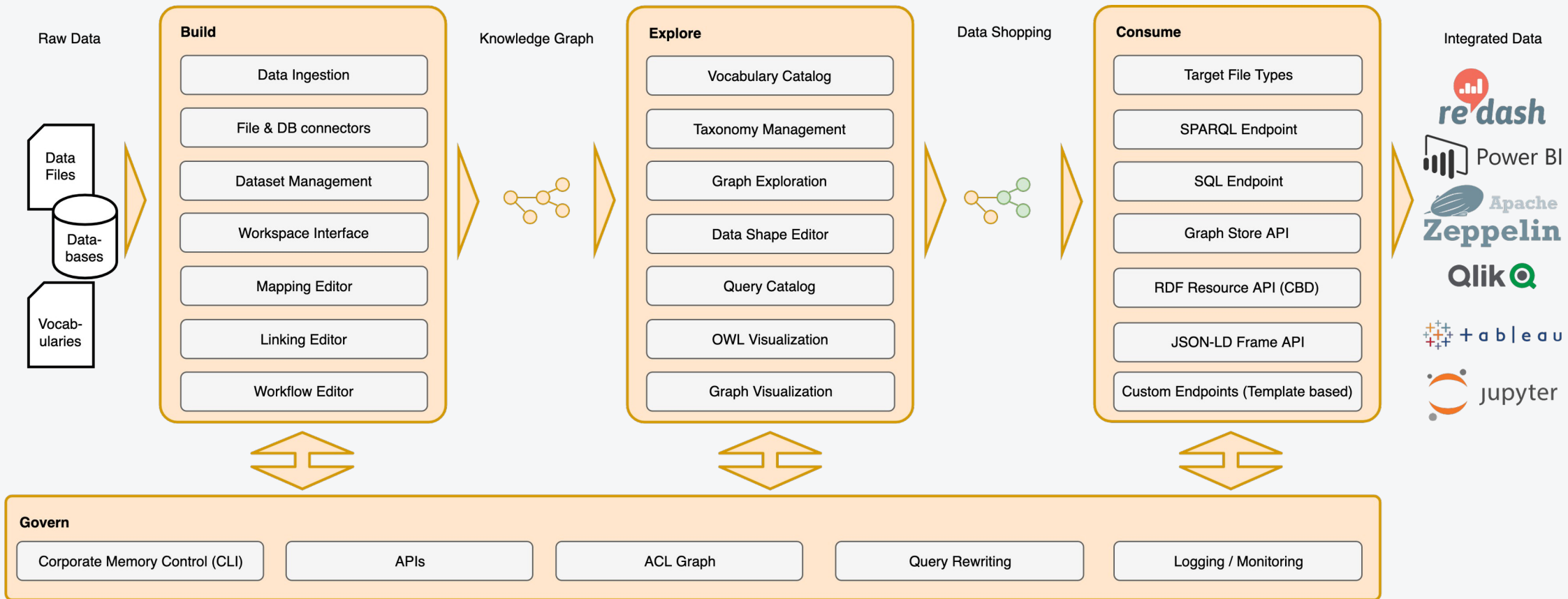
🗑️

```
graph LR; S1[Source path: gn:alternateName] --> SE1[String equality: Threshold 0.0, Weight 1]; T1[Target path: label] --> SE1; S2[Source path: gn:countryCode] --> SE2[String equality: Threshold 0.0, Weight 1]; T2[Target path: coy.code-3166-2] --> RE[Regex extract: (.*)-.*]; RE --> SE2; SE1 --> AND[And: Weight 1, Aggregation]; SE2 --> AND;
```

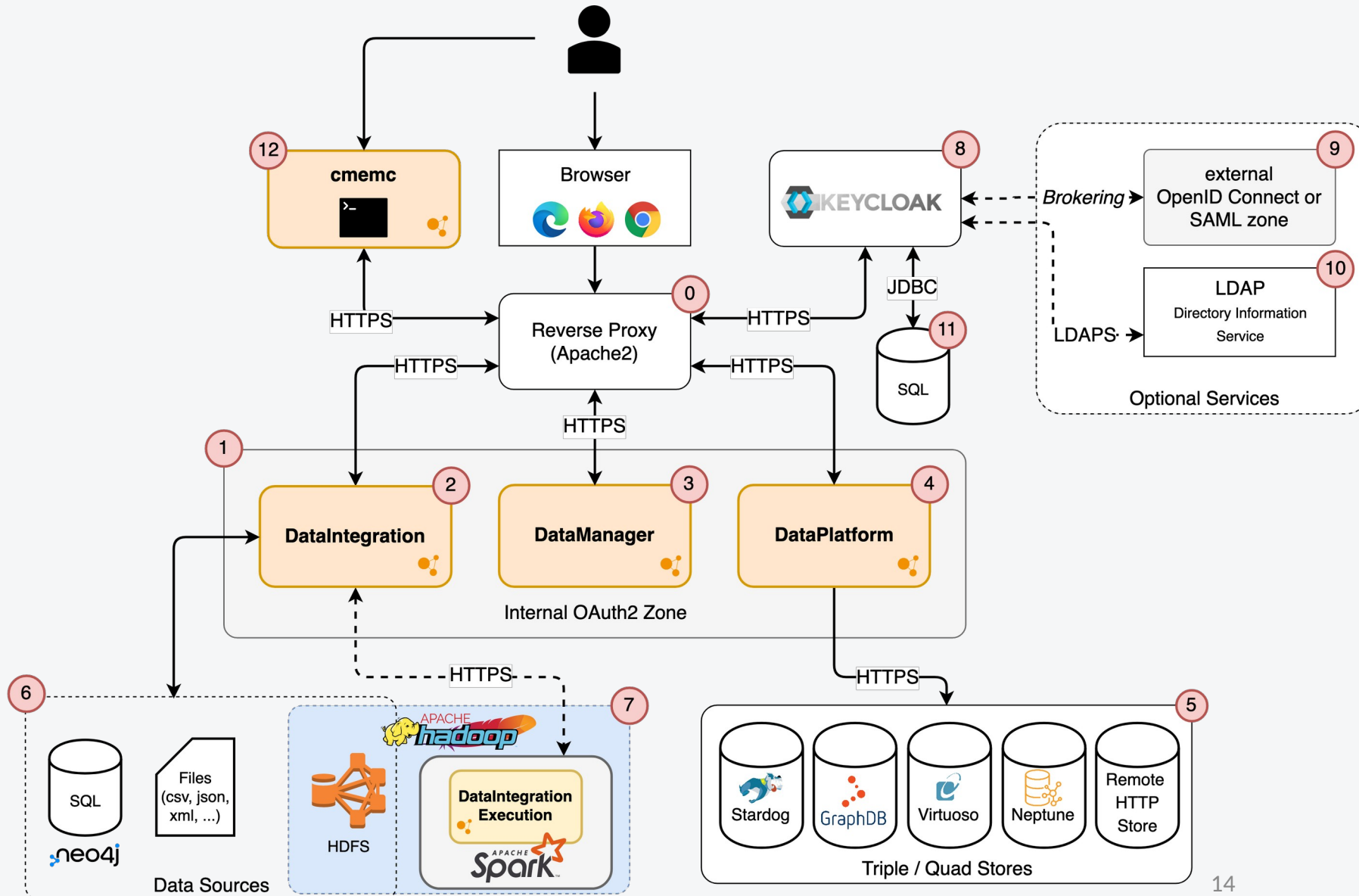
The Mapping Workflow



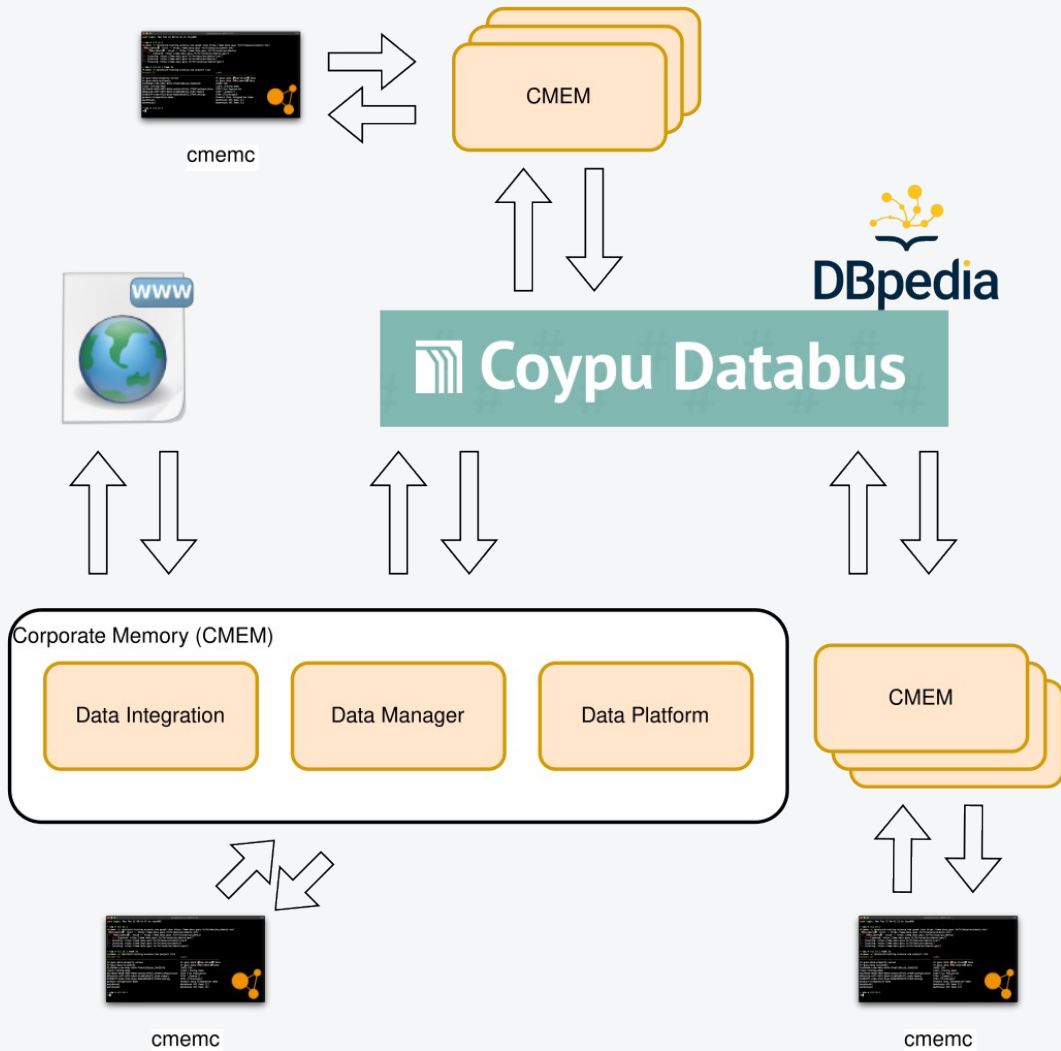
User Journey and Functional Areas



Reference Architecture Corporate Memory



Integration with the Databus



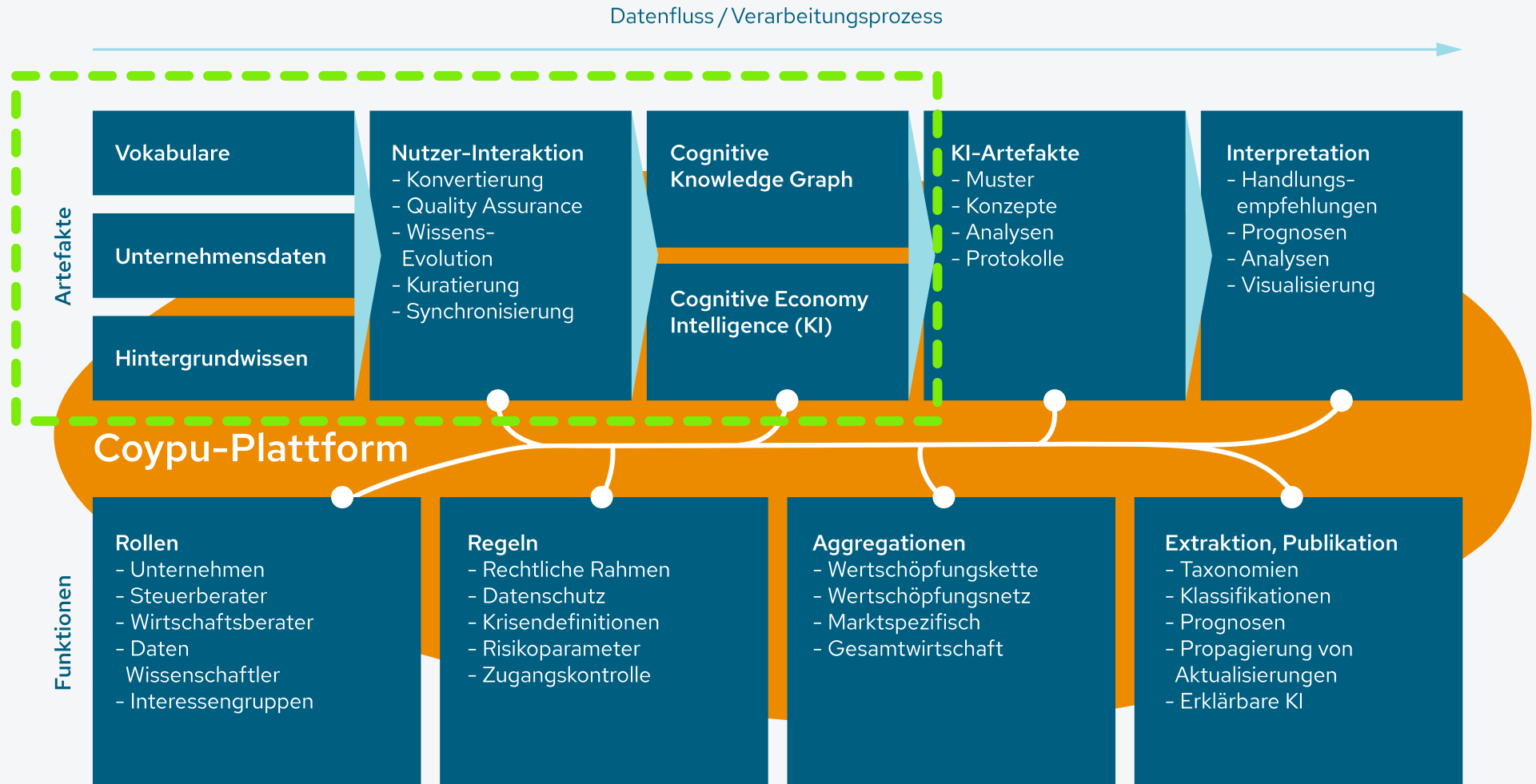
The screenshot shows the Coypu Databus web interface with the following sections:

- Header:** Coypu Databus, OPEN BETA, Login, SparQL Endpoint, API Documentation.
- RECENT ACTIVITY:** A line chart showing 'Uploaded Data (Gbyte)' over time. The data shows a sharp increase in June.
- SEARCH:** A search bar with the text 'Search the Databus...' and navigation icons.
- NEW ARTIFACTS:** A section featuring an artifact titled 'Implisense Company Data Dump' by user 'nsteinert', with a description 'Test dump of full Implisense company data' and a timestamp '2022-06-10T13:16:45.257Z'.
- MOST ACTIVE USERS:** A table listing active users and their data contributions.

User	Uploads	Derived Data
nsteinert	1	107 KB

DBpedia
Global and Unified Access to Knowledge
Graphs
© Copyright 2019 by DBpedia. All Rights Reserved.

The Coypu-Plattform





eccenca GmbH
Hainstraße 8
D-04109 Leipzig
Germany

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



eccenca
mastering complexity

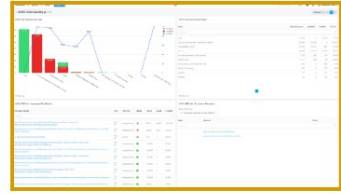
Hans-Chr. Brockmann
Geschäftsführer

Hainstraße 8
D-04109 Leipzig
Germany

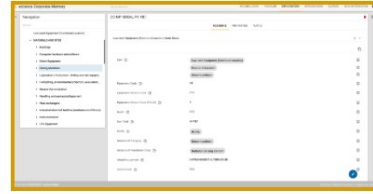
+49 511 3365 2810
+49 173 3698 610

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

BACKUP

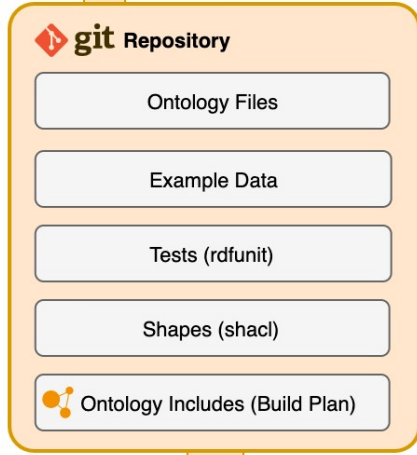
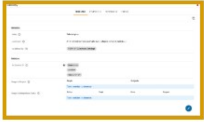


Data Quality Dashboards

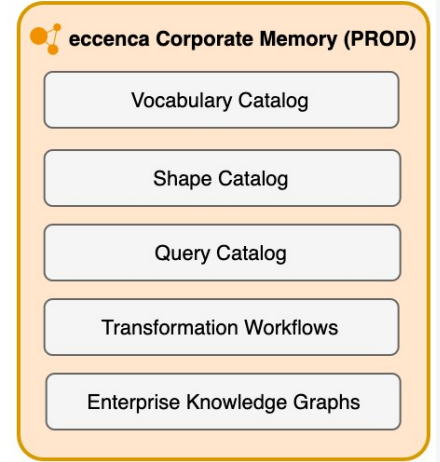
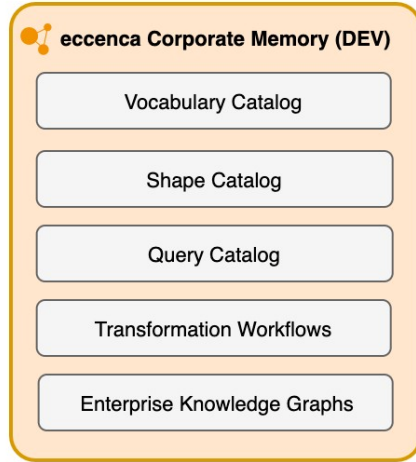


Data Quality Drilldown

Git Working Policy
(discusses branches, releases, ...)



cmemc



"Going Live"

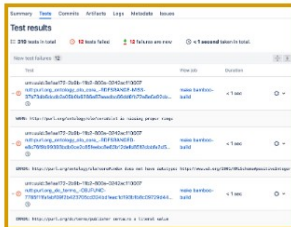
Triggered Builds

Build Server

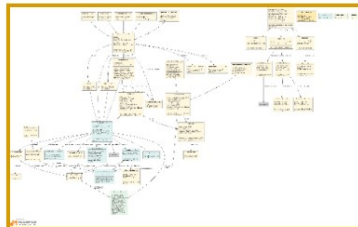
Ontology Build Plan
(Testing, Reasoning, Artifact Creation)

Test Live-Data
Deploy on Success
Annotate Resources

Build Artifacts



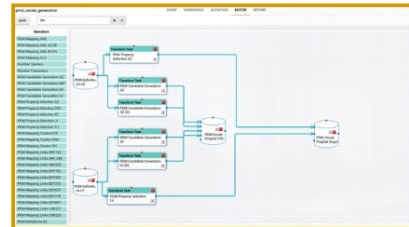
Error Report



Static Visualization



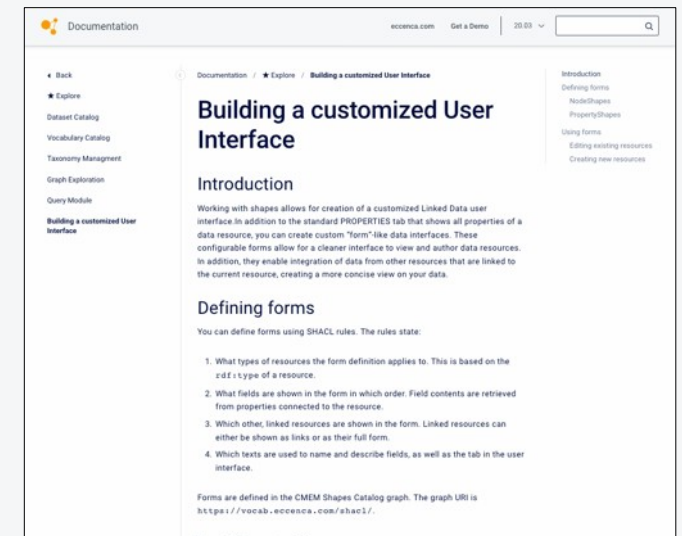
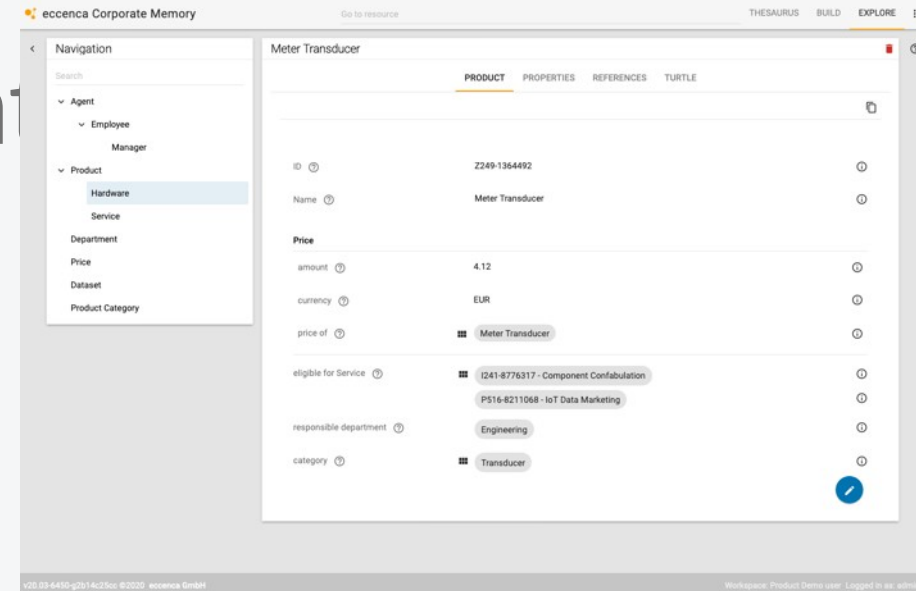
Widoco
Documentation



Ontology Reuse Package

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>

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

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

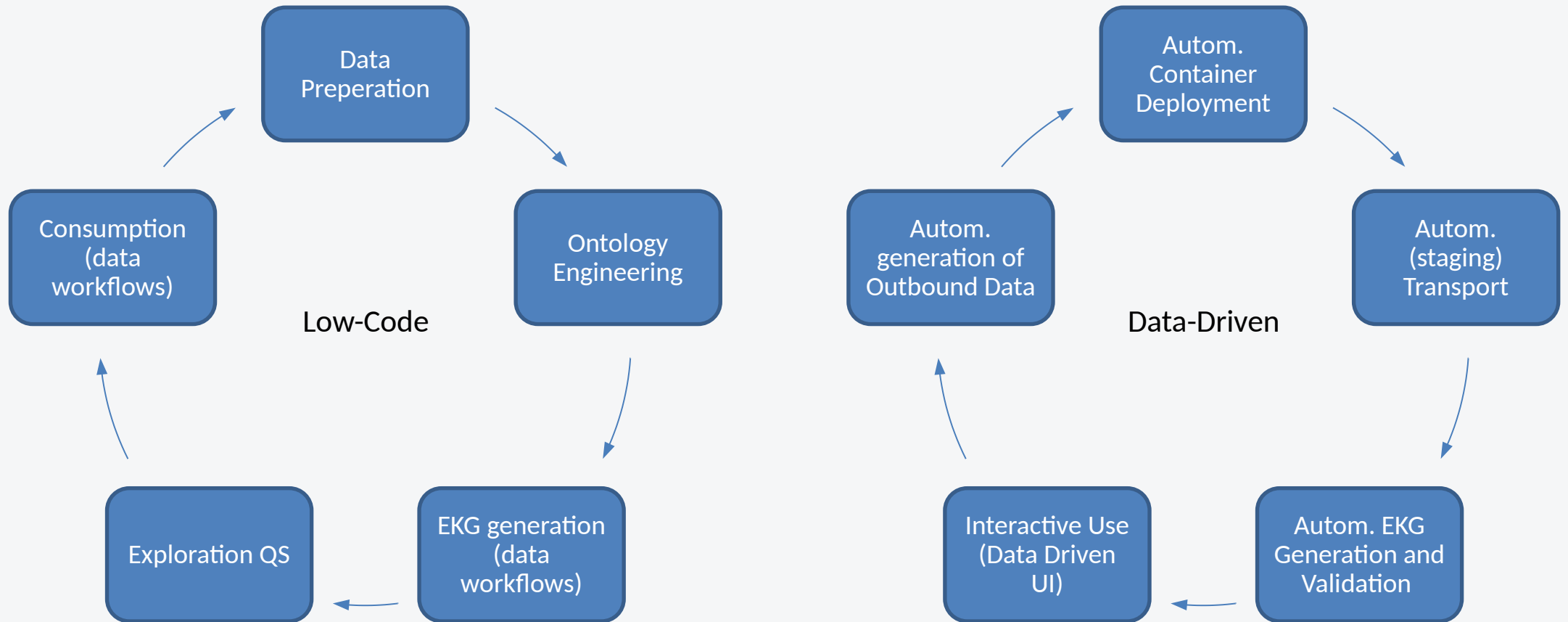
DevOps based Ontology Engineering Pipeline

- We apply code and testing principles to RDF datasets, esp. ontologies Engineering
- We use git to manage revisions of our ontologies and bamboo as the DevOps tool
- Any Ontology tool / IDE possible that consumes/produces RDF
- Automatic test generation and quality checks using builds with each commit
- cmemc is used to automate the workflow between ontology engineering, corporate memory development and instance provisioning

-

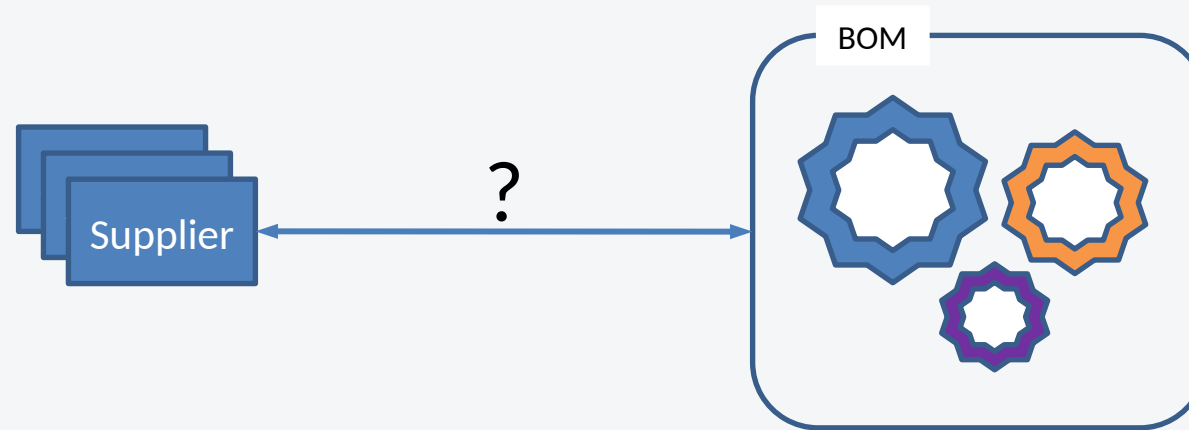
<https://bamboo.eccenca.com/browse/SCHEMA-ECCDSM-99/test>

eccenca Build- and Run-Cycles



SAMPLE USE CASES

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



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

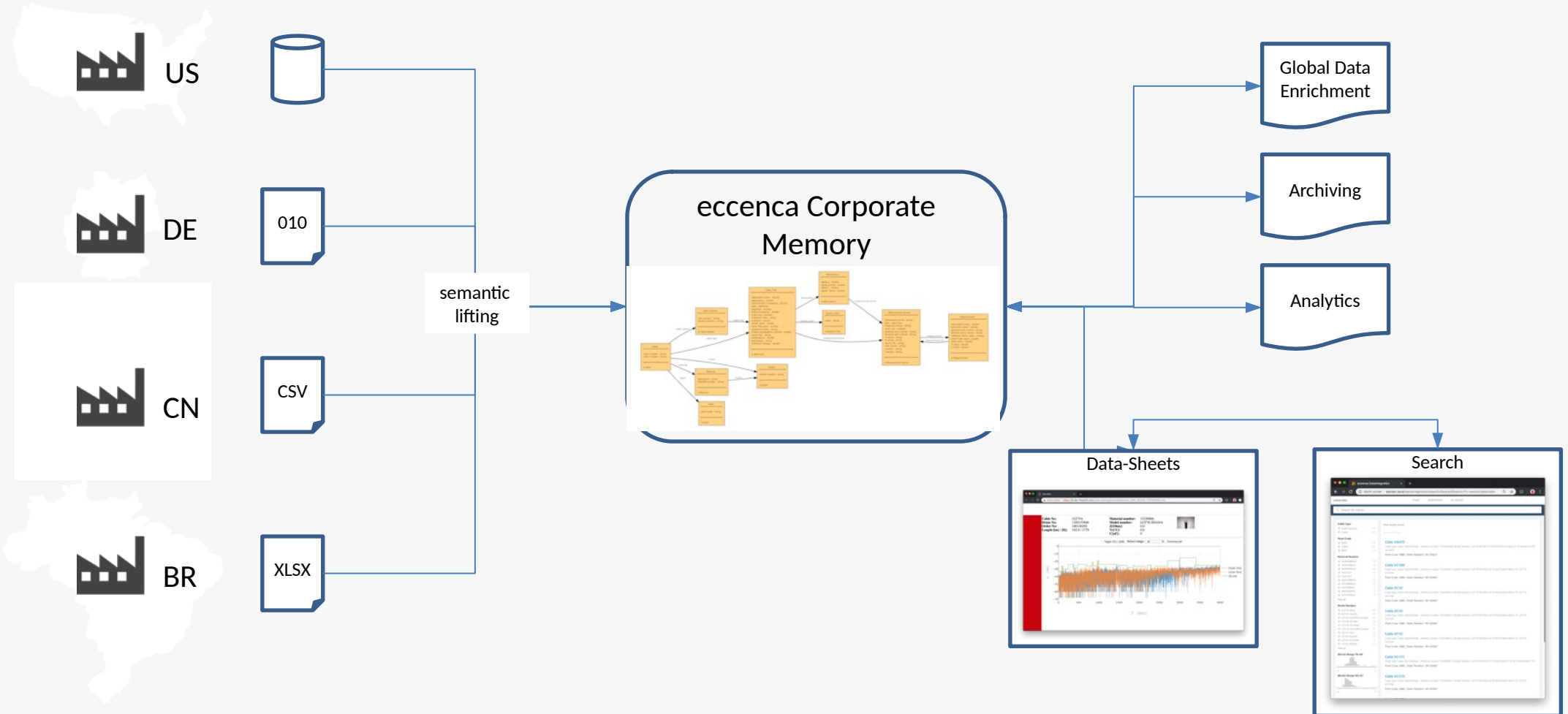
MPN	Supplier
293D105X9016A2##E3	VISHAY
B45196E3105K10	KEMET
T491A105K016AS	KEMET
T491A105K016AT	KEMET
TAJA105K016RNJ	AVX

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



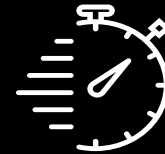
Project Benefits

- 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

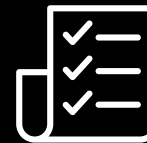
BENEFIT Cases



- 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

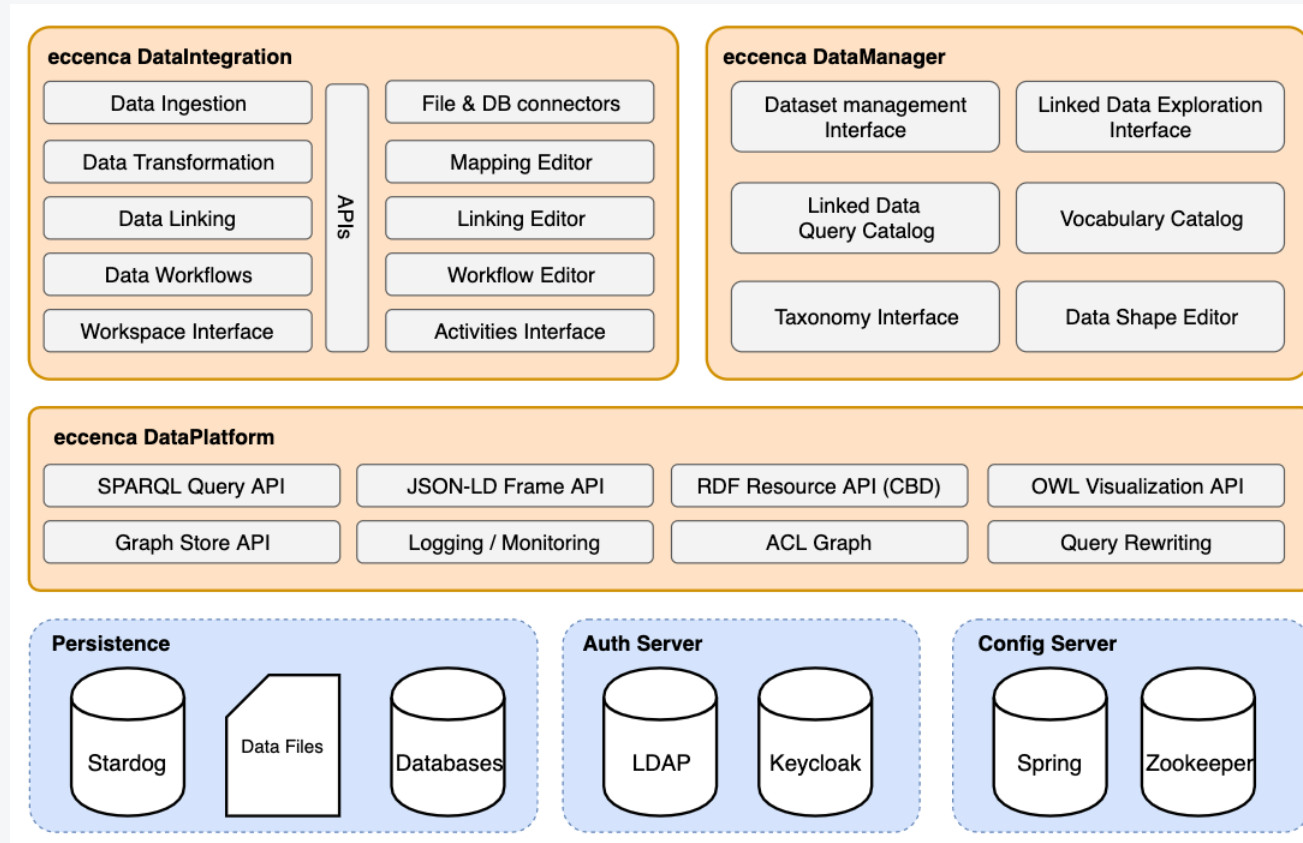
Initiator: DBPedia, Linked GeoData etc.
Betreiber: Datenportal der EU-Kommission
Team-Leitung: Prof. Dr. Sören Auer

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

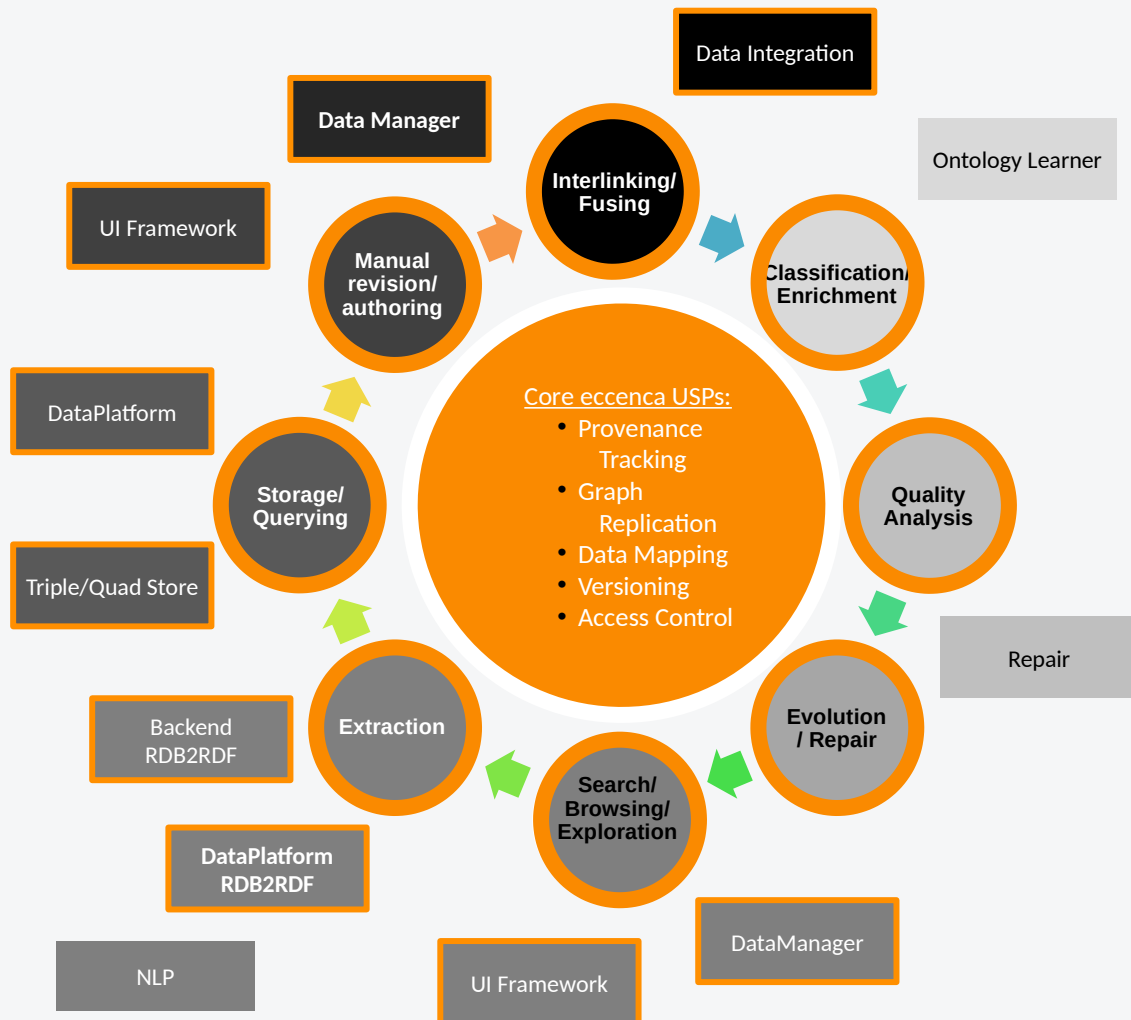
What

- § 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

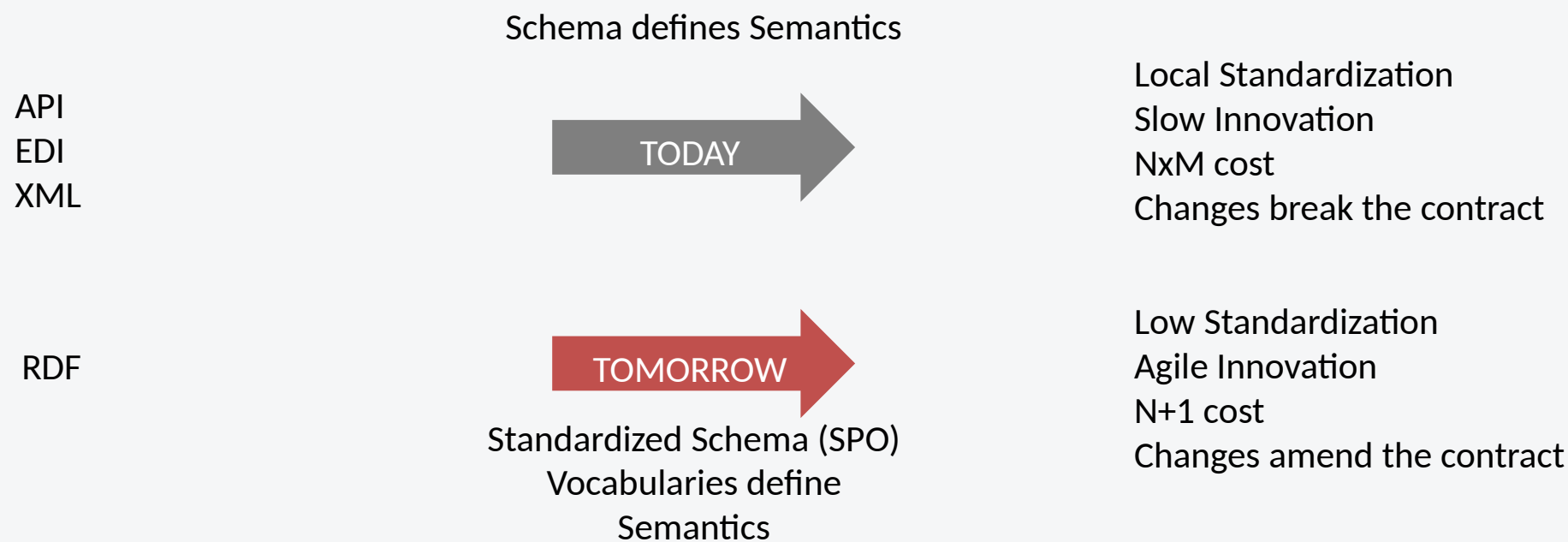
Linked Data Life Cycle



- 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



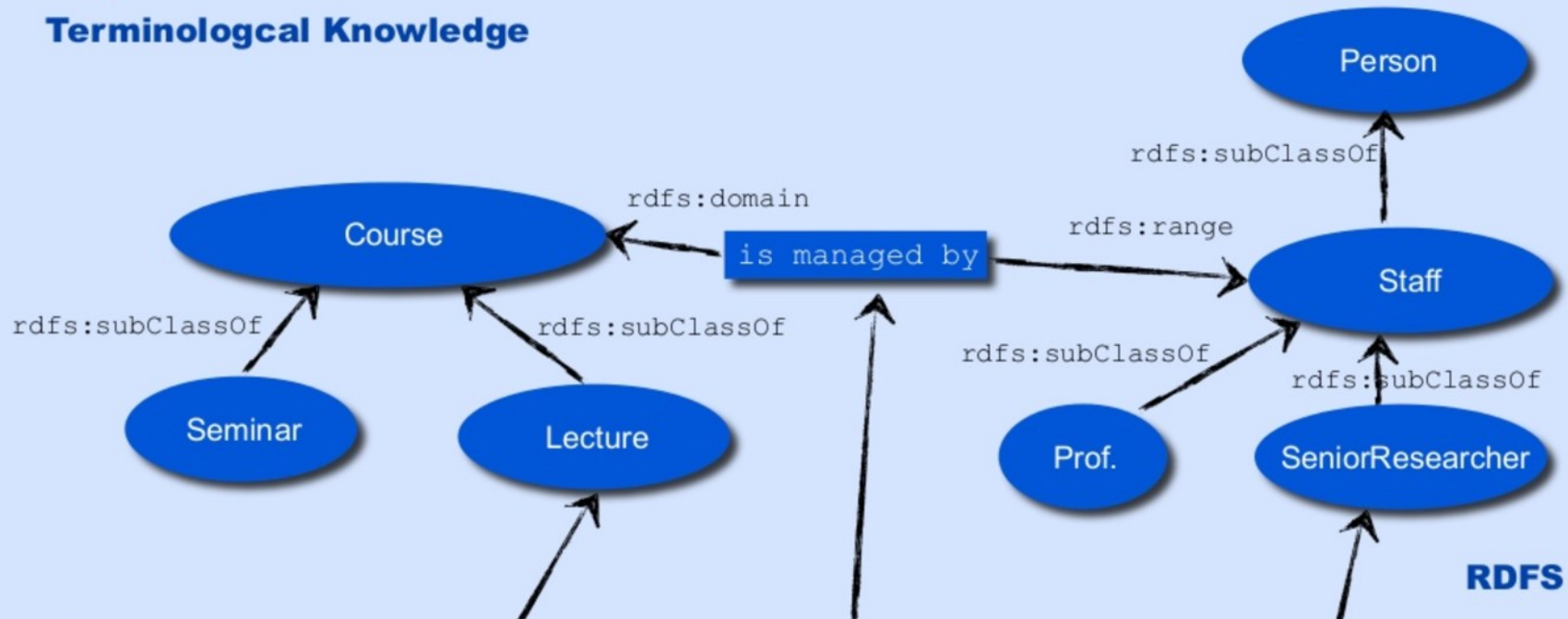
Other Use Cases

- 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.

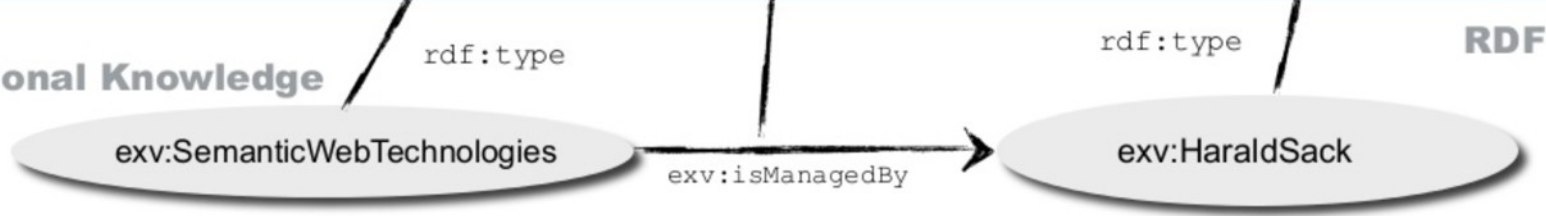
RDF + RDFS

13

T-Box Terminological Knowledge

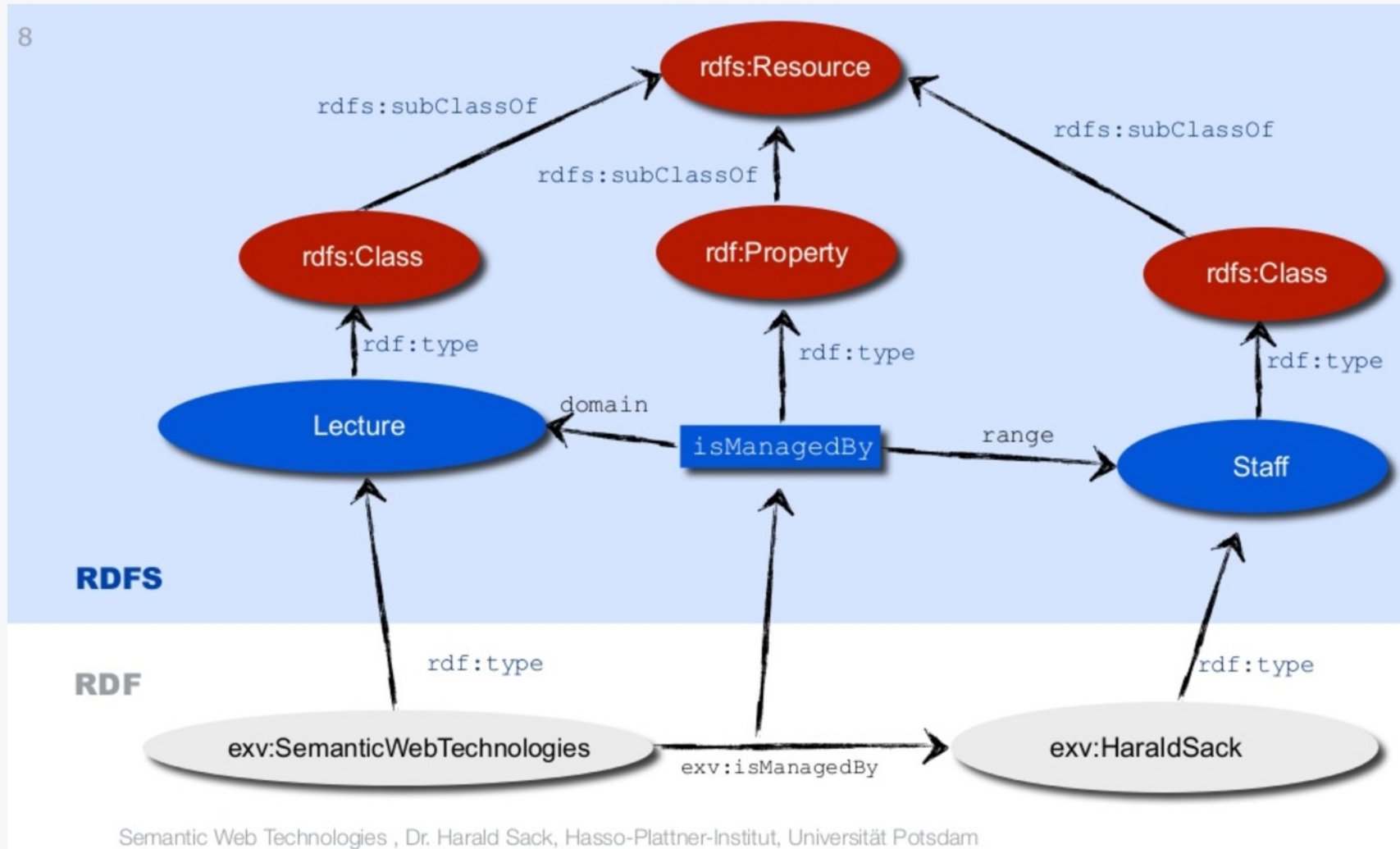


A-Box Assertional Knowledge

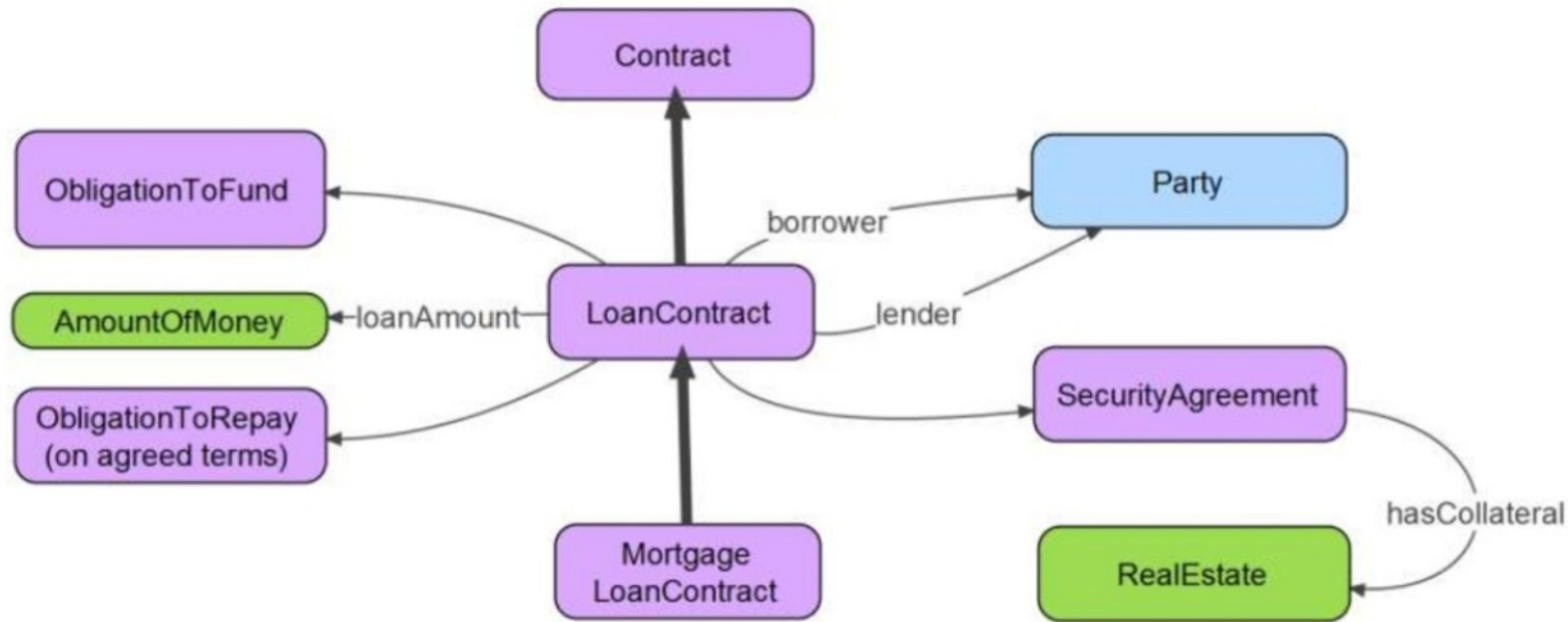


Classes, Properties, Instances

8



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 .
```

