



## Recommender Systems based on Knowledge Graphs

EBISS (European Big Data Management & Analytics Summer School)

Albin Ahmeti Researcher

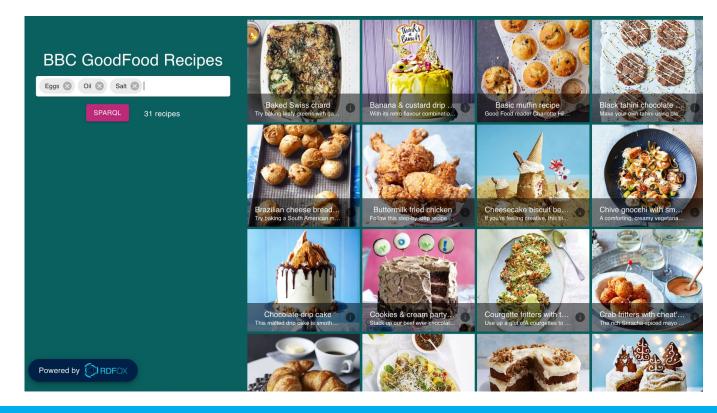




- **Part I:** The ingredients and their origins
- Part II: The recipe
- Part III: The secrets of cooking
- **Part IV:** The proof is in the pudding! (Food recommender demo)

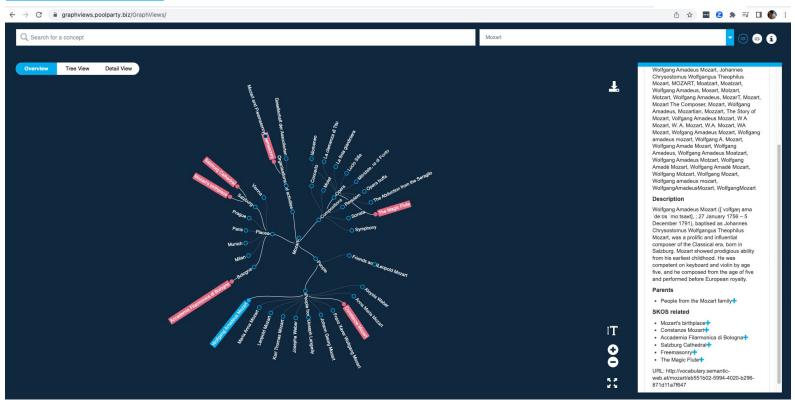
## Part I: The ingredients and their origins





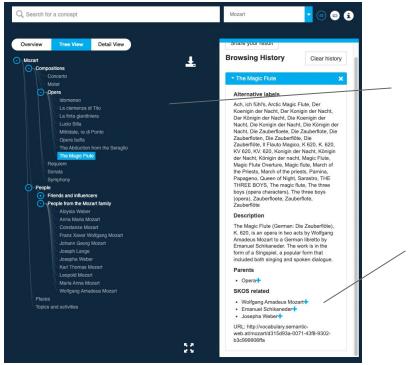
© Semantic Web Company





### https://graphviews.poolparty.biz/GraphViews/



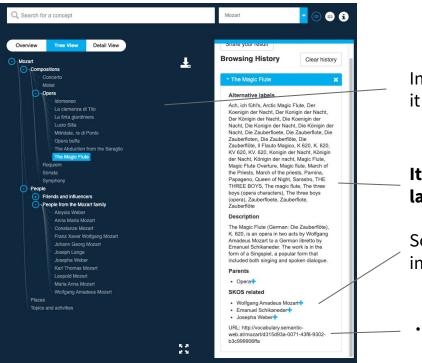


### In a different 'Tree view' you actually see that it is actually organized in a hierarchy.

# Some parts can have 'graph-like' relations instead of hierarchical tree.

### https://graphviews.poolparty.biz/GraphViews/





In a different 'Tree view' you actually see that it is actually organized in a hierarchy.

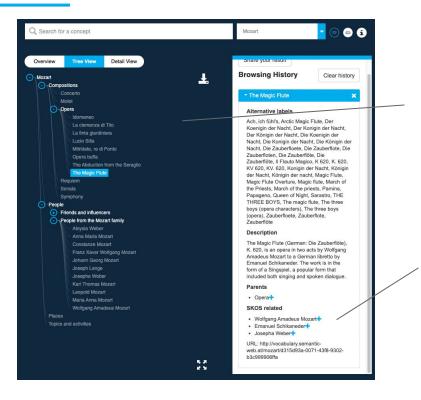
# It captures knowledge in different languages.

Some parts can have 'graph-like' relations instead of hierarchical tree.

...pay attention: one URI.

### https://graphviews.poolparty.biz/GraphViews/





In a different 'Tree view' you actually see that it is actually organized in a hierarchy.

Some parts can have 'graph-like' relations instead of hierarchical tree.

Jumping from [Mozart] to [The Magic Flute], and then to [Emanuel Schikaneder] or [Josepha Weber] and so on..sounds familiar?

```
https://graphviews.poolparty.biz/GraphViews/
```

# Any of you use Obsidian?



Thoughts

Images 2023-05-23 2023-05-24 2023-05-30

Answer Set Programming?

Collaborative filtering

CWA vs OWA

Data integration - NF crossovers

Datalog rules

DBpedia linking

ElasticSearch recommender

Enterprise Knowledge Graphs

Explain command

### **Enterprise Knowledge Graphs**

EKGs = Taxonomies + Ontologies + Instance data. Combination of virtual graphs + materialized graphs. You can create virtual graphs in GraphDB via Ontop.

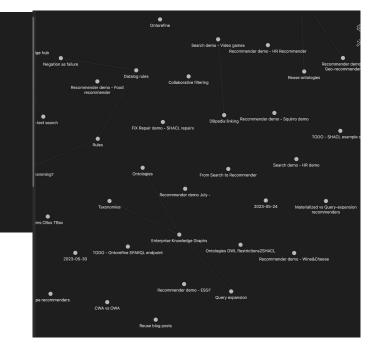
[[Taxonomies]][[Ontologies]][[Query expansion]]

### Materialized vs Query-expansion recommenders

Make a slide where you group recommenders in two groups: - Materialization-based - Query-rewrite / Query-expansion

In this case: • Geo-recommender is materialized. • Wine & Cheese is query expansion-based.

materialization #query-rewriting



## Zettelkasten method

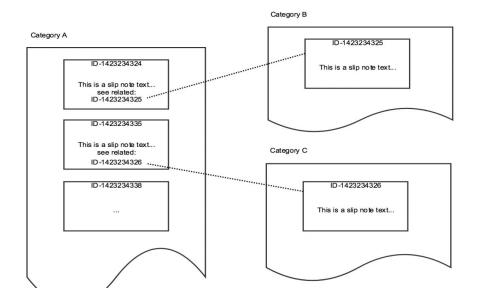


- Obsidian uses '**Zettelkasten method**' by Niklas Luhmann to organize knowledge.
- You create your ideas/notes/excerpts and then you **tag** them, e.g. #query-rewriting, originally referred to as 'index cards'.
- You also **link** to all the relevant notes that you can think of, e.g. [[Taxonomies]].
- It becomes a powerful system to store, search, retrieve and to serendipitously discover a new piece of knowledge that you didn't know existed before, or you couldn't think of.

## Zettelkasten method (cnt.)



• Index cards with pointers (links) grouped into Categories.



## Zettelkasten method (cnt.)



- From the perspective of search you can either:
  - Follow the tags, such as #query-rewriting
  - Follow the links, such as [[Taxonomies]]

## Follow the tags

- Think of 'Faceted Search based on concept'
- The tag or keyword is a concept

## • Follow the **links**

- Think of 'Similarity'
- The link is another item that is similar to the current by some *measure*.

## Zettelkasten method within Obsidian



Q tag:#materialization Aa ⊗ 🗄	← → Materialized vs Query-expansion recommenders []	D :
4 results 💩 File name (A to Z) 👻		
V Enterprise Knowledge Graphs	Materialized vs Query-expansion	
#materialization #virtual-graphs	recommenders	
<ul> <li>Materialized vs Query-expansion</li> <li>recommenders</li> </ul>	Make a slide where you group recommenders in two groups:	
#materialization #query-rewriting	Materialization-based	
Recommender demo - Geo-recommender 1	Query-rewrite / Query-expansion	
#materialization	In this case:	
Search demo - HR demo 1	<ul> <li>Geo-recommender is materialized.</li> </ul>	
#materialization	<ul> <li>Wine &amp; Cheese is query expansion-based.</li> </ul>	

	Enterprise Knowledge Gr × +	~	<i>ଊ</i> , ଊ, # ∷≡	
Q tag:#materialization Aa 🛛 🗄	$\leftarrow  ightarrow$ Enterprise Knowledge Graphs			
4 results  File name (A to Z)			materialization	4
✓ Enterprise Knowledge Graphs	Enterprise Knowledge Graphs			2
#materialization #virtual-graphs			datalog	1
	EKGs = Taxonomies + Ontologies + Instance data.		dbpedia	1
<ul> <li>Materialized vs Query-expansion</li> <li>recommenders</li> </ul>			geonames	1
	Combination of virtual graphs + materialized graphs.		KGs	1
#materialization #query-rewriting			llms	1
Recommender demo - Geo-recommender	You can create virtual graphs in GraphDB via Ontop.		modelling	1
	·····		obdm	1
#materialization	[[Taxonomies]][[Ontologies]][[Query expansion]]		ontologies	
Search demo - HR demo			overarching	1
#materialization	#materialization #virtual-graphs		query-rewriting	1

1. Find all notes with tag #materialization 2. Jump to the notes that contain the tag #virtual-graphs

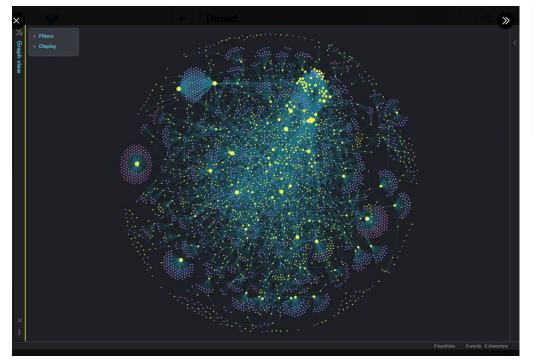
2'. Jump to the notes (subset) that contain both the tags including #virtual-graphs

2". Jump to [[Taxonomies]], [[Ontologies]] or [[Query Expansion]]

## Graph view of linked notes in Obsidian



...



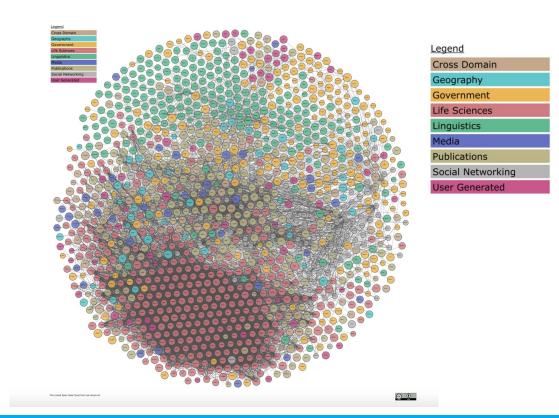
Obsidian @obsdmd Massive graph of @myownslave , 3,639 notes: 1:07 am · 30 Sep 2020

9 Retweets 1 Quote 81 Likes

## Linked Open Data cloud



- Stats (as of May 2020):
  - 1,255 datasets
  - 16,174 links
- DBpedia at the center as
   LOD nucleus with links to other datasets.
- Data transformed to RDF using ETL (Extract Transform Load) and mappings.





- Knowledge Graphs (KGs) describe a domain of interest and typically consist of:
  - Conceptual or abstract representation of knowledge in the form of ontologies (TBox);
  - Individuals or **instance data (ABox)**;
  - Controlled vocabularies in the form of **taxonomies (CBox)**.

In the classical sense of AI:

```
Knowledge base = TBox + ABox.
```

In our setting:

KG = TBox + ABox + (CBox).

## Minimal KG example

### • TBox:

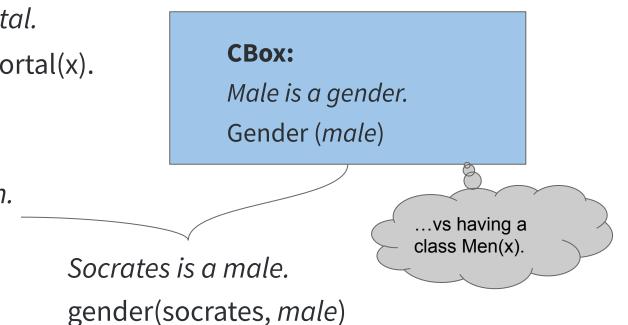
All persons are mortal.

 $\forall x \operatorname{Person}(x) \Rightarrow \operatorname{Mortal}(x).$ 

## • ABox:

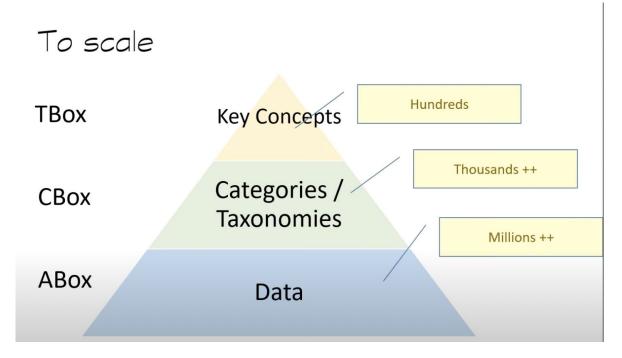
*Socrates is a person.* Person (socrates).





## The Pyramid representation of a KG...in numbers





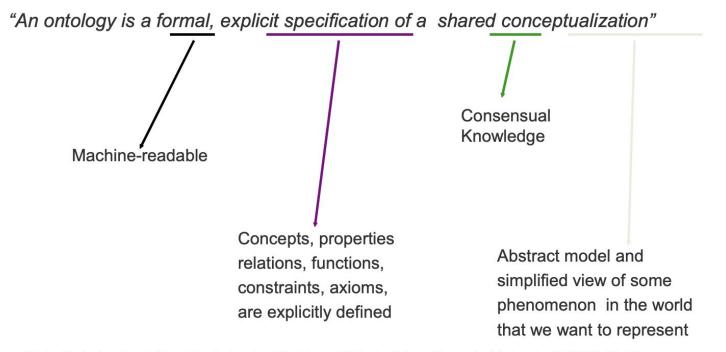
Each of them are **managed** separately and they are of different **size**.

This delineation including CBox, is also due to **governance**.

The image taken from the gist <u>ontology</u>.

## **Ontology definition**





Studer, Benjamins, Fensel. Knowledge Engineering: Principles and Methods. Data and Knowledge Engineering. 25 (1998) 161-197

## Main purposes of ontologies (TBox)



- Use it as a "traditional" *schema*...to **describe the data** in an expressive way with classes and relations
- Schema mapping
  - Map (semi-)structured data e.g. relational databases with <u>different schemas</u> to an <u>unified schema</u>;
  - Use case of heterogeneous data integration for (Enterprise) Knowledge Graphs
    - Also virtualized approach to data integration, so called

### "Ontology-based data access".

- ...to **infer implicit facts** based on explicit facts and ontology statements (so-called *axioms*), typically done by the *reasoner* as part of a triple store.
  - The inference is done by the reasoner recursively until no new implicit triples are inferred, i.e. when we reach "a fixed point."

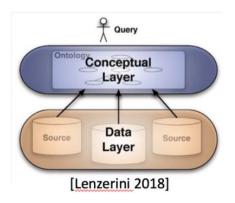
## **Ontology-based Data Management (OBDM)**



- Conceptual layer is represented by an ontology, i.e., encodes the semantics;
- Query answering via *query rewriting* techniques for exposing data via mappings Ontology-based Data Access (OBDA);
- Incremental adding of sources to the unified ontological schema;
- Updates are less explored [Ahmeti, PhD thesis 2020].

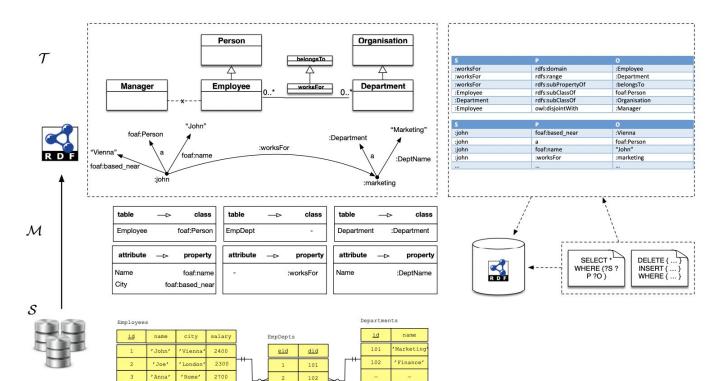
 $\langle \mathcal{S}, \mathcal{M}, \mathcal{T} 
angle^{\mathbb{S}, \mathbb{L}}$ 

- Semantics of querying and update
- Query and update language



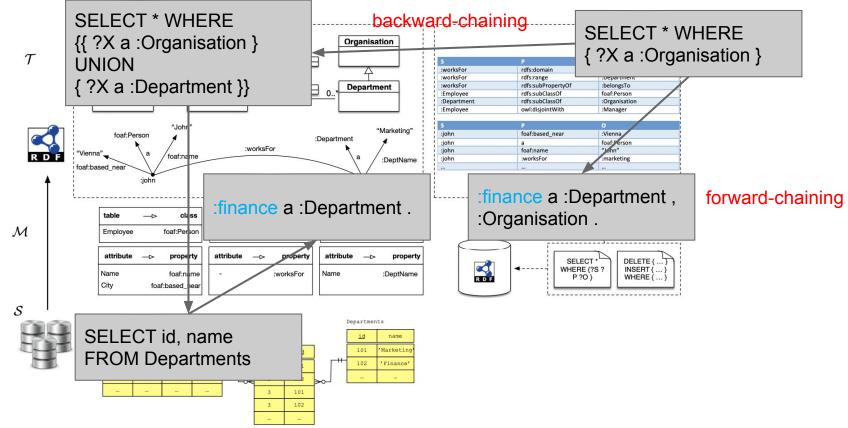
## **Ontology-based Data Access (OBDA)**





## **Ontology-based Data Access (OBDA)**





## DBpedia example





	S	P	0	
Ī	:hasMother	rdfs:subPropertyOf	:hasParent	1
	:DauphinOfFrance	rdfs:subClassOf	:RulerOfFrance	
	:Place	owl:disjointWith	:Organisation	
	:marie	:hasMother	:maria_t	
	:marie	:hasSpouse	:louis	
	:maria_t	а	:RulerOfAustria	
	:louis	а	:DauphinOfFrance	
{{Templa   mapToO   mappin {{Propert {Property	yMapping   template Mapping   template	Property = name   onto	m   ontologyProperty = royalAn logyProperty = foaf:name }}	ithem }}
{{Pronert	vManning   template	Property = title   ontol	ogyProperty = title }}	

1

{Pr {{PropertyMapping | templateProperty = title | ontologyProperty = title }} {{DateIntervalMapping

| templateProperty = reign

-

| startDateOntologyProperty = activeYearsStartYear

| endDateOntologyProperty = activeYearsEndYear }}

....

{Infobox royalty	· · · · · · · · · · · · · · · · · · ·	tones most, + +
consort=yes	+ directoralized	falogʻulofitaris, knoivetu 🧧 🤁 🔍
Image - Louise Elisabeth Vigles-Laroum - Marie Antoinette de la Bose - Oogle Art Project pg capton - "Marke Antoinet with the Resc"-Vignet tab (Tables Elisabeth Vigles Le Brun)), 1783. Isaccession - Elisabeth Commerci of Farces and Neurori] (ingle - 1 Ale particle - Adjamenter 1791) (ingle - 1 Ale particle - Adjamenter 1792) (ingle - Laroum - Market - Adjamenter 1792) (ingle - Laroum - Market - Adjamenter 1792) (ingle - Laroum - Market - Adjamenter 1792) (ingle - Clare - Farces) (ingle - Clare - Farces) (ingle - Clare	Exercise Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Construction Constr	Arrise Tan Marce Antoinette Arris Marca and Arrise any agent Arris Marca and Arrise Arrise and Arrise Arrise March Antoinette Arrise and Arrise Arrise Arrise March Antoinette Arrise and Arrise Arrise Arrise March Arrise an Arrise Arrise and Arrise Arrise and March Arrise and Arrise Arrise and Arrise Arrise March Arrise and Arrise Arrise and Arrise Arrise March Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise March Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise March Arrise Arrise Arrise Arrise Arrise Arrise Arrise March Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arrise Arri
[Incose : [Incose] + Habburg-Loranite]] [Incose : [Incose] + Kinosen : Engenzon [Incose : [Incose] + Kinosen : Engenzon [Incose] + Kinosen : [Incosent] + Kinosen :	Newsition New Williams Commission Recent charges Recent charges This Marcines have Recent charges Recent charge	Emprove Frances 1: In April 1771, que te ter namingar to Local Augusto, Da Inservar Anton Frances, Die ausward filt eille Ga- Samarra Anton Frances, Die ausward filt eille Ga- attalisage aufbalter Laus XI in Aber (1971), Aber aussi auf alle geschlichte Laus XI in Aber, Straft Eille Barnes, Fil- Isteller, Hennestry für processing auf balaus), bei filt bestellter Barnes (1970), and and and and and and Australia Barnestry auf alle Saugerite Barn Franken auf- Barnest annesse, auf aufbalaus, Straft eil Barnes, Filteren Barnes, sollter Barnessen, Straft er Barnes Hennes annesse, aufbalaus Aller, Alle Aber Hennes annesse, aufbalaus Aller, Alle Aber Hennessen Aller aufbalaus Aller, Alle Aber Hennessen Aller aufbalaus Aller Aller Aber Marchannessen Process Theorem Barnesi Allere Alleres Barnessen Alleres, annessen Alleres Alleres Alleres Barnessen Alleres, annesse Alleres Alleres Alleres Barnessen Alleres, annessen Alleres Alleres Alleres Barnessen Alleres, annesse Alleres Alleres Alleres Barnessen Alleres, annesse Alleres Alleres Alleres Barnessen Alleres, annessen Alleres Alleres Alleres Alleres Alleres Alleres Alleres Alleres Alleres, annessen Alleres Al

004440-0 H



 $\mathcal{M}$ 

 $\mathcal{T}$ 

 $\mathcal{S}$ 

## **Ontology expressivity in PoolParty**



- Ontology expressivity support in PoolParty within OWL 2 QL
  - "Minimal" RDFS (no axiomatic triples) + some constructs/axioms of OWL
    - It turns out that this is approx. the OWL fragment that is used in the wild
  - Minimal RDFS: rdfs:subClassOf, rdfs:subPropertyOf, rdfs:domain, rdfs:range
  - OWL: *owl:unionOf*, *owl:FunctionalProperty*, *owl:SymmetricProperty*, *owl:inverseOf*, *owl:disjointWith*.
- Closed (CWA) vs Open World Assumption (OWA)
  - **Restrictions:** *owl:Restriction, owl:someValuesFrom, owl:allValuesFrom, owl:minCardinality, owl:maxCardinality, owl:hasValue* 
    - Captured in SHACL using CWA

## **Functionality in PoolParty**



- Functionality: In PoolParty implemented using CWA
  - E.g. [:manages a owl:FunctionalProperty].
  - It doesn't allow you to add more than one object or value for that property/attribute.
- If directly done on a triple store (via INSERT DATA), you will get a semantics mismatch due to CWA vs OWA
  - o [:a :manages :b . :a :manages :c .] -> [:b owl:sameAs :c]

Note that in SHACL processor you would get "maxCardinality" violation.





In OWA the following stands:

```
[:john a skos:Concept; :livesIn :Vienna . :livesIn
rdfs:domain :Person] -> [:john a :Person]
```

In PP (CWA) the following stands:
 [:john a skos:Concept; :livesIn :Vienna . :livesIn
 rdfs:domain :Person] -/> [:john a :Person]
 [:john a skos:Concept; a :Person] ->\* [:john :livesIn XYZ]

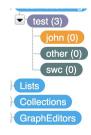
```
-> infer
```

- -/> does not infer
- ->\* possibly infers (constraint check)

## RDFS (cnt.)



- Subclass + Subproperty reasoning in Thesaurus Manager
- Note: [Employee -> Person -> Thing] [worksFor -> belongsTo]



0	<u>an.poolparty.biz/</u> e, Person, Tl	<u>TestReasoning/2</u> hing	
Details	Notes	Documents	Linked Data
SKOS	• TestR	easoningScheme	<u>•</u> +
belongsTo (	i)		
⊗ <u>other</u> ⊗ <u>swc</u> ⊘			
worksFor (i)	)		

## RDFS (cnt.)



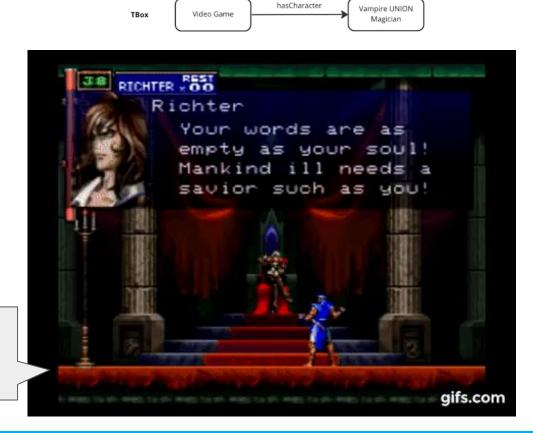
- Domain and range violations can not be repaired, but only removed
- Otherwise, ontology domain/range definitions have to be changed if possible

Semantics_2021     Equipment (125)     EQUIP-A7137 (0)     EQUIP-A9812 (0)	EQUIP-A7137 <i>https://poc.poolarity.biz/Equipment/ID/EQUIP-A7137</i>	+ Add to Collection 🚫 Add to Blacklist 🚫 Add to ExactMatch 👕	Delete Concept
- EQUIP-C1298 (0) - EQUIP-C3331 (0)	Details Notes Documents Linked Data Triples	Visualization Quality Management History	
EQUIP-E2129 (0) EQUIP-E2246 (0)	SKOS <b>ISO14424_Scheme</b> A • +		
EQUIP-E2246 (0)	hasComponentMaintainableItem ④	Additional information (i)	Апу Туре
	${}^{\oslash}$	Ð	
EQUIP-E6672 (0)	hasFailure 🚯	Additional information in free text as applicable (i)	Any Type
EQUIP-E8991 (0) EQUIP-E8992 (0)	0	Ø /image/sample_image.png ⊕	xsd:string
EQUIP-E9762 (0) EQUIP-F3231 (0)	hasMaintenance 🚯	0	
EQUIP-F3238 (0)	Ø	Agitator Speeds (RPMs)	Integer
EQUIP-F3331 (0) EQUIP-F3332 (0)	hasPart (i)	⊗ 160.0 ▲	
EQUIP-F4152 (0)	Ø	Amps (i)	Float
EQUIP-F7137 (0) EQUIP-F7138 (0)	hasSubunit ③	⊙ 1.5 <b>▲</b>	
	Ø	Attachment Speeds (RPMs) (1)	Integer
EQUIP-H2247 (0)		⊗ 110.0 ▲	

## A few words on class disjointness



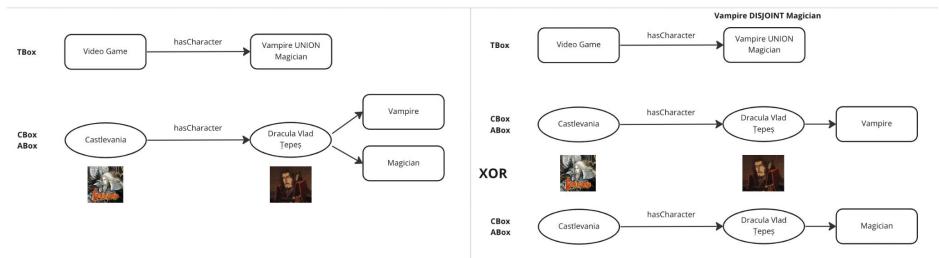
Hint on how to beat the game: Strip power from the Dracula by introducing the disjoint constraint in the TBox...



Vampire DISJOINT Magician

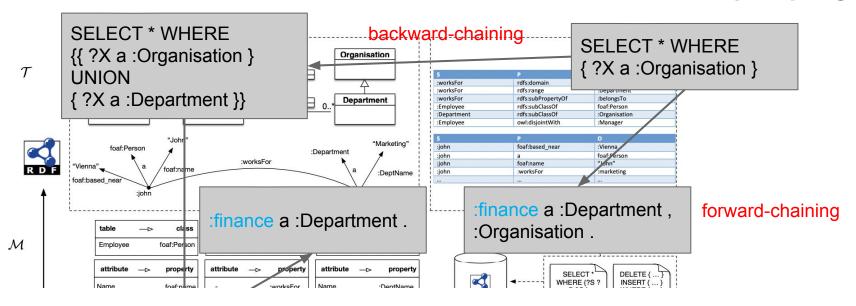
## A few words on class disjointness (cnt.)





- In data integration setting, we have two models that satisfy the constraints in TBox and data in ABox/CBox = KG,
  - $I \models TBox, I \models ABox:$ 
    - I1={ :CastlevaniaSoN :hasCharacter :DraculaVlad . :DraculaVlad a :Vampire}
    - I<sub>2</sub>={ :CastlevaniaSoN :hasCharacter :DraculaVlad . :DraculaVlad a :Magician}

## Query rewriting vs materialization in PoolParty



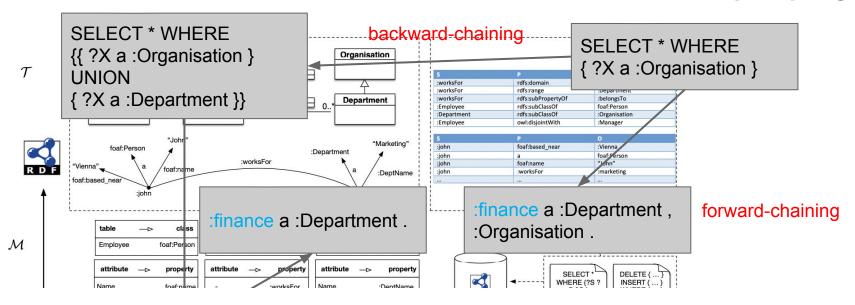
#### Proposition

Given a triple store  $G = \mathcal{T}_{\mathcal{P}} \cup \mathcal{A}$ , where  $\mathcal{T}_{\mathcal{P}}$  is a PoolParty ontology (TBox) and  $\mathcal{A}$  is instance data (ABox). Let  $\mathcal{A}_{imp}$  be all the triples that can be inferred from  $\mathcal{A}$  and  $\mathcal{T}_{\mathcal{P}}$  via forward-chaining. Then the following holds for a SPARQL query q, and the rewritten one q' formed from q and  $\mathcal{T}_{\mathcal{P}}$  via backward-chaining

 $q(x,\mathcal{A}\cup\mathcal{A}_{imp})=q'(x,\mathcal{A}).$ 

poolparty.

## Query rewriting vs materialization in PoolParty



#### Proposition

...the same equation holds even for virtualized instance data  $\mathcal{A}$  and q'.

 $q(x,\mathcal{A}\cup\mathcal{A}_{imp})=q'(x,\mathcal{A}).$ 

poolparty<sub>®</sub>

## GraphDB as an OBDA system





The default option to store, i.e persist the data.

Ontology-based data management aka *virtualization*.

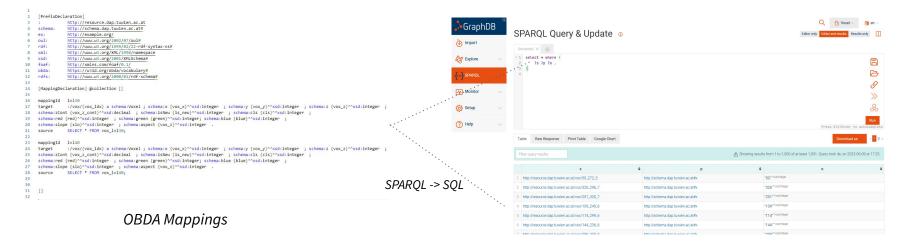
#### Edit Repository: Voxel

Cancel

Location	Local		\$
Repository ID*	Voxel		Ø
Repository description			
Ontop configuration			
Database driver	Generic JDBC Driver	\$	Test connection
JDBC properties file*	voxel.properties	Ø	1 Upload file
OBDA or R2RML file*	voxel-mappings.obda	Ø	1 Upload file
Ontology file			1 Upload file
Constraint file			1 Upload file
DB metadata file			1 Upload file
- Destant secondary			

## GraphDB as an OBDA system (cnt.)





jdbc.url=jdbc:sqlite:/Users/ahmetia/Documents/HopsVoxels/vo

x VIE 2022W data.sqlite

jdbc.driver=org.sqlite.JDBC

jdbc.user=

jdbc.password=

### JDBC Properties

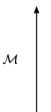
## Map relational databases to an unified schema



Ehttps://sharepoint-demo/hr/Employee/l> <https://sharepoint-demo/hr/emplo> "I"^xsd:long; <https://sharepoint-demo/hr/age "39"^xsd:long; <https://sharepoint-demo/hr/age "39"^xsd:long; <https://sharepoint-demo/hr/jobTitle> "Research Assistant"; <https://sharepoint-demo/hr/jobTitle> "Research Assistant"; <https://sharepoint-demo/hr/link-to-cv> "B0,000.00"; <https://sharepoint-demo/hr/link-to-cv> "https://docs.google.com/document/d/ 1v90505VdrUmB9W897vX1U66fDXWBrsqbTXhcjjB0A/edit"; <https://sharepoint-demo.com/hr/images "https://drive.google.com/open?id=10T54B9WgRIFN3a92CzpaB3uKyCVJIaTD"; <http://sharepoint-demo.com/hr/images "https://drive.google.com/open?id=10T54B9WgRIFN3a92CzpaB3uKyCVJIaTD"; <https://sharepoint-demo/hr/jbloy2/dd/ddt";</pre>









					HR database	
Employee ID	Name	Age	Job Title	Gross Salary	Link to CV	Pic
1	Adah Punjabi	39	Research Assistant	80,000	https://docs.google.com/document/d/1v9Q505VC4rUm9VW897vX41U6ofDXVBrsqblXhCj80A/edit	https://drive.go
2	Charles Roberts	29	Machine Learning Expert	134,000	https://docs.google.com/document/d/1nvqZpJJYa4i48nUuVBGBSqLNN_5wD24dNb_zdNNJ9E/edit	https://drive.go
3	Martina Folanini	49	Business Analyst	80,000	https://docs.google.com/document/d/1vr56Fwcl_Usu8GPqDYJUom5/OddTXbsidq1BtYvndis/edit#	https://drive.go
4	Mohamed Azure	35	Research Assistant	70,000	https://docs.google.com/document/d/1044fTl8kthymdOmkt90cSoFxszA5oPpl/s5xY80UPt6g/edit	https://drive.go
5	Roland Jones	30	Senior Software Developer	120,000	https://docs.google.com/document/d/18LFcb4od0VOIIwbUG-ite-FBg-8dwgiOXrJk4i5MJNo/edit	https://drive.go
6	Tom Brian	44	Software Developer	95,000	https://docs.google.com/document/d/1HKX2DKw-MENG9K_wpzK0gTywcptPinnMt2XYJaXCuq0/edit	https://drive.go
7	Winston Mckenzie	62	Artificial Intelligence Expert	150,000	https://docs.google.com/document/d/1sr3cW_xiztbag1_BmpSME3VrAdbBEgFIBeK4ryzwEE/edit	https://drive.go
8	Xiaoxi Chen	48	Senior Software Developer	115,000	https://docs.google.com/document/d/18gWW8dJulLm9bn6AZtnM4AWXyPmWVUY01xU2-bEncwL/edit	https://drive.go
9	Albert Sandler	35	Research Assistant	80,000	https://drive.google.com/open?id=1EmawZofVNzD1TvP51t-wY1BlvK5e9volb9Vy_Y6zE8	https://drive.go
10	Alexadra Otero	41	Machine Learning Expert	134,000	https://drive.google.com/open?id=1dCrWbys5RQluXYCIRsZoq4iu0NXIRXizAWt05KEQQU	https://drive.go
11	Alicia Prochesko	44	Business Analyst	80,000	https://drive.google.com/open?id=1YthvGk3EY8sgGGZ_n2NqxVUlqUIB9PBM02I4WgInh5M	https://drive.go
12	Alina Dawson	39	Research Assistant	70,000	https://drive.google.com/open?id=1uRBjbXgL9VQQV0uAALFMk8gluWr29aksFc8iu-jTRMg	https://drive.go
13	Amalia Chater	54	Senior Software Developer	120,000	https://drive.google.com/open?id=1Xwx9rVjQflhXtrGJ-HVZ83gxP4_WTTPO1wBjpgXQI4	https://drive.go

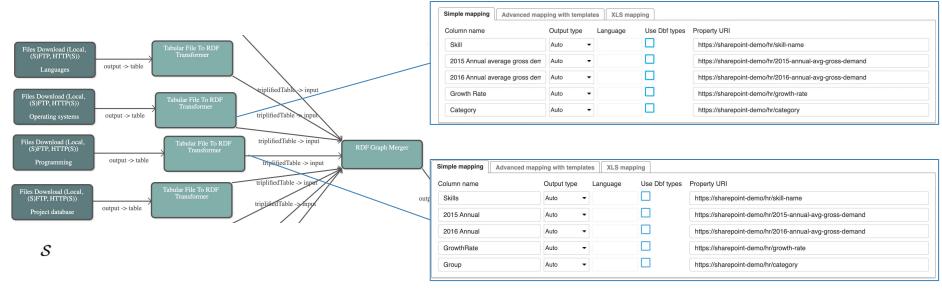
# Map relational databases to an unified schema (cnt.) poolparty.



-> fetches data from both sources.

© Semantic Web Company

# Map relational databases to an unified schema (cnt.) poolparty.



 $\mathcal{M}$ 

INSERT DATA { :skill1 <<u>https://sharepoint-demo/hr/categor</u>> :Programming}

-> Not clear in which data source to propagate the update? Column Category or Group? **View update** problem. [NP-complete problem]

© Semantic Web Company

# **Taxonomy definition**

#### What is a taxonomy?

*Controlled* and *organized* 

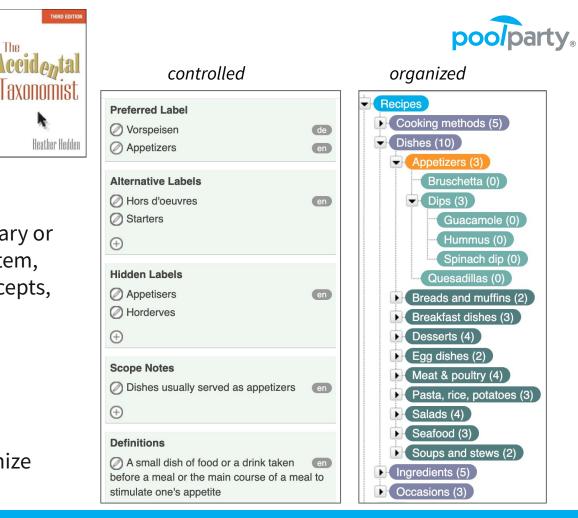
#### 1. Controlled:

A kind of controlled vocabulary or knowledge organization system, based on unambiguous concepts, not just words: things, not strings

The

#### 2. Organized:

Concepts are arranged in a structure of hierarchies, categories, or facets to organize them.



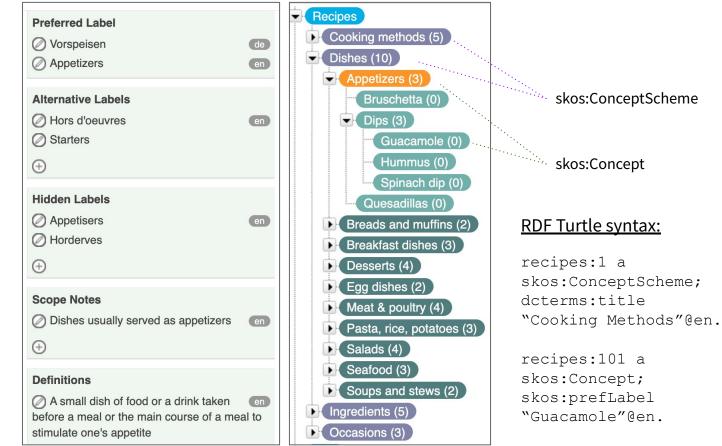




- SKOS is a standard ontology endorsed by W3C that is used to represent thesauri, taxonomies and controlled vocabularies.
- SKOS ontology contains of set of **classes**, **relations** and **attributes** that are used to describe a taxonomy.
- This ensures "no vendor lock-in" meaning that a taxonomy created from one system should be easily exported and imported to another system, maintaining the **exact semantics**.

### poolparty.

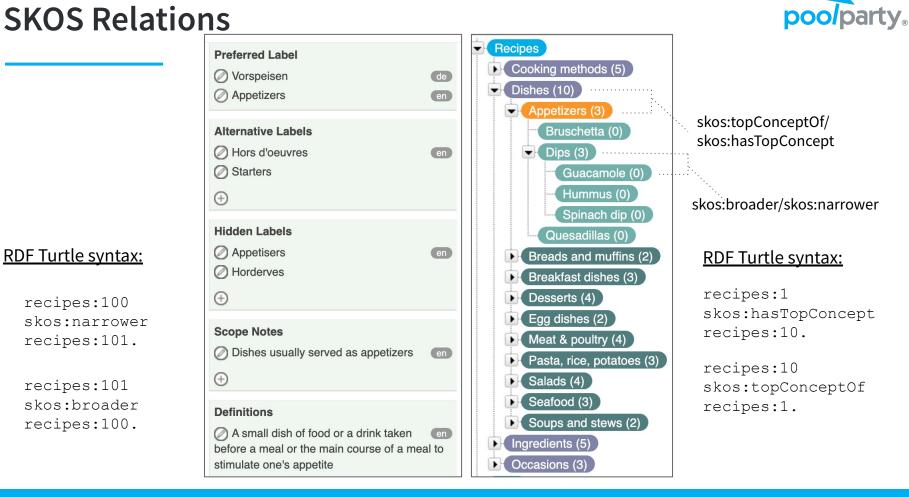
#### **SKOS Classes**



#### **RDF** Turtle syntax:

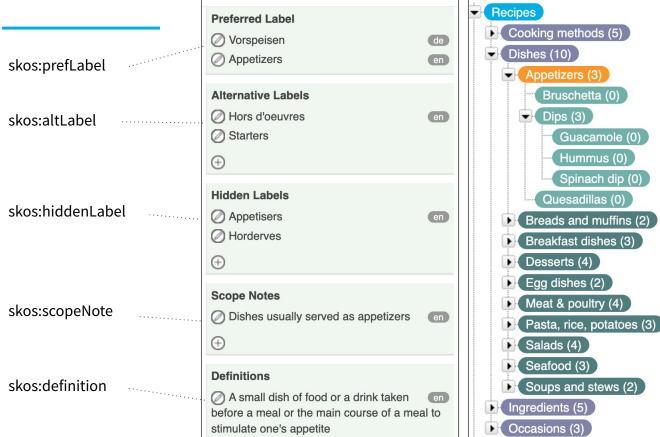
recipes:100 skos:narrower recipes:101.

recipes:101 skos:broader recipes:100.



#### SKOS Attributes





and few more: skos:notation skos:example

. . .

...

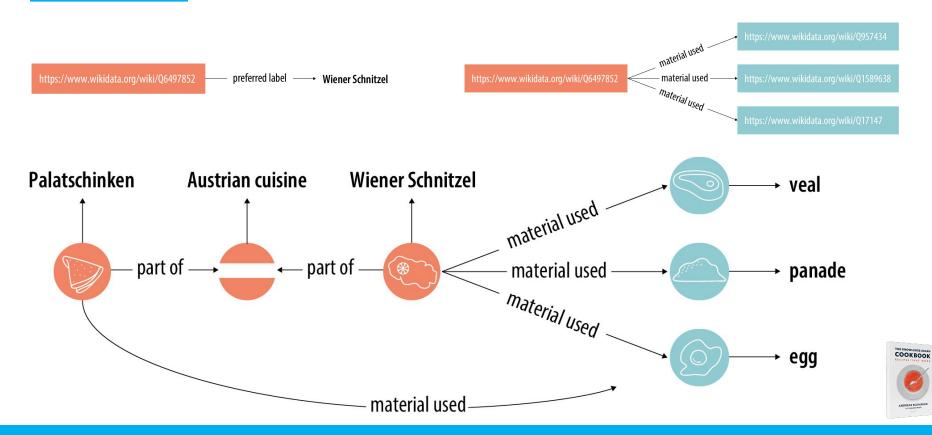
#### **RDF Turtle syntax:**

recipes:15
skos:prefLabel
"Appetizers"@en,
"Vorspeisen"@de;
skos:altLabel "Hors
d'oeuvres"@en,
"Starters"@en

© Semantic Web Company

# Extending concepts with relations





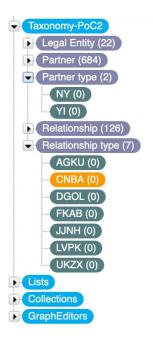
# Extracting taxonomies (KGs) from tabular data



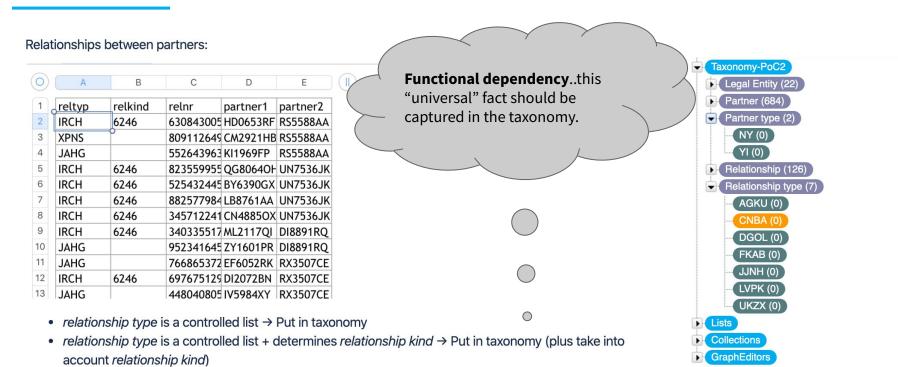
#### Relationships between partners:

	А	В	С	D	E
-					
	reltyp	relkind	relnr	partner1	partner2
	IRCH	6246	630843005	HD0653RF	RS5588AA
1	XPNS	0	809112649	CM2921HB	RS5588AA
[,	JAHG		552643963	KI1969FP	RS5588AA
	IRCH	6246	823559955	QG80640F	UN7536JK
	IRCH	6246	525432445	BY6390GX	UN7536JK
	IRCH	6246	882577984	LB8761AA	UN7536JK
[	IRCH	6246	345712241	CN4885OX	UN7536JK
[	IRCH	6246	340335517	ML2117QI	DI8891RQ
	JAHG		952341645	ZY1601PR	DI8891RQ
,	JAHG		766865372	EF6052RK	RX3507CE
	IRCH	6246	697675129	DI2072BN	RX3507CE
Ī.	JAHG		448040805	IV5984XY	RX3507CE

- relationship type is a controlled list → Put in taxonomy
- relationship type is a controlled list + determines relationship kind → Put in taxonomy (plus take into account relationship kind)
- partner1 + partner2 → Put in taxonomy

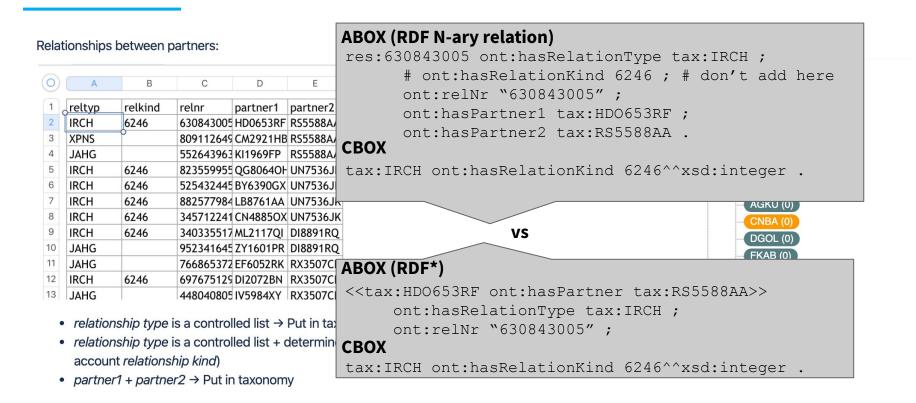


# Extracting taxonomies (KGs) from tabular data (cnt.) poolparty.

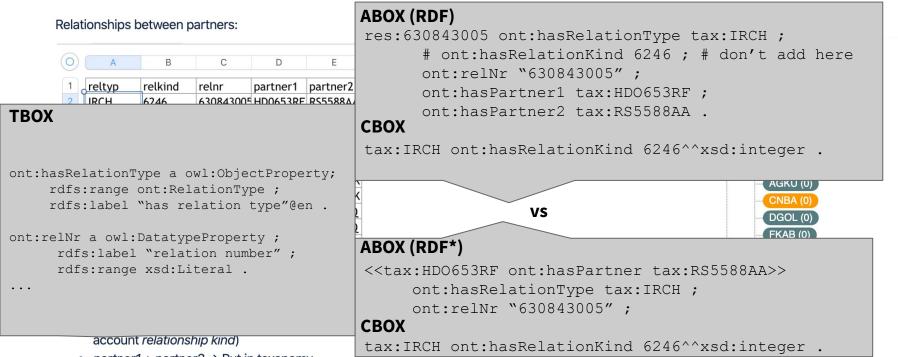


• partner1 + partner2 → Put in taxonomy

# Extracting taxonomies (KGs) from tabular data (cnt.) poolparty.



# Extracting taxonomies (KGs) from tabular data (cnt.) poolparty.



• partner1 + partner2  $\rightarrow$  Put in taxonomy

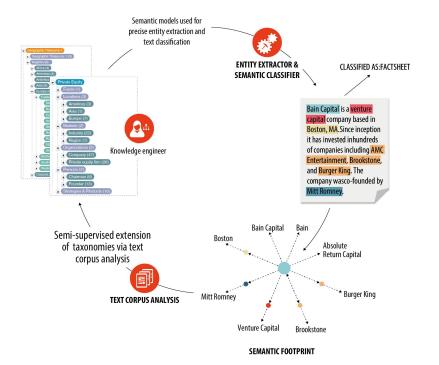


- Organize knowledge in a **hierarchical (SKOS) and graph-based** (exploiting OWL relations) using concepts with a well-defined meaning.
- Use it for taxonomy-based information retrieval, powering semantic search and recommender systems.
- Tag documents, text and other **unstructured data** with concepts in a given language, i.e., **entity extraction** leveraging NLP techniques (stemming, tokenization).
  - The document is tagged against the skos:prefLabel, skos:altLabel and skos:hiddenLabel.

#### • Text position, frequency will have an impact on the scoring of concepts.

### High-level view of entity extraction



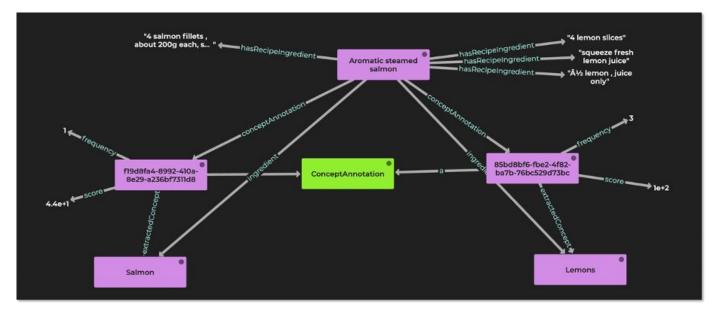




#### **Entity extraction**

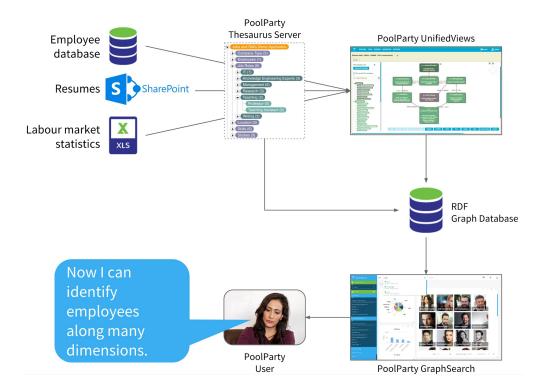


#### • Extraction of taxonomy concepts (CBox) from instance data (ABox).



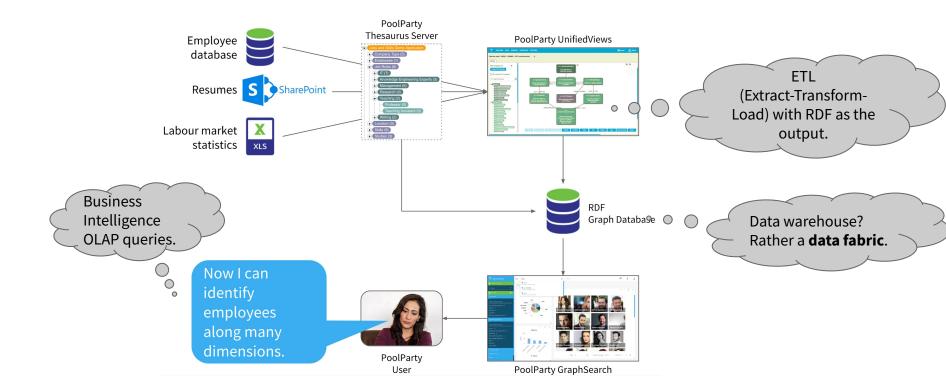
#### Use case: Graph-based HR Analytics





#### Use case: Graph-based HR Analytics





#### **Graph-based HR Analytics: Datasets**



A	В	С	D	E	F	
				0	HR database	
Employee ID	Name	Age	Job Title	Gross Salary p	Link to CV	Pic
1	Adah Punjabi	39	Research Assistant	80,000	https://docs.google.com/document/d/1v9Q505VC4rUm9VW897vX41U6ofDXVBrsqblXhCjj80A/edit	https://drive.goo
2	Charles Roberts	29	Machine Learning Expert	134,000	https://docs.google.com/document/d/1nvqZpJJYa4I48nUuVBGB5qLNN_5wD24dNb_zdNNJ9iE/edit	https://drive.goo
3	Martina Foianini	49	Business Analyst	80,000	https://docs.google.com/document/d/1vr56Fwcl_Usu8GPqDYJUom5VOddTXbsldq1BtYvndis/editil	https://drive.goo
4	Mohamed Azure	35	Research Assistant	70,000	https://docs.google.com/document/d/1044fTi8kthymdOmkt90cSoFxszA5oPpVs5xYB0UPt6g/edit	https://drive.goo
5	Roland Jones	30	Senior Software Developer	120,000	https://docs.google.com/document/d/18LFcb4od0VOIIwbUG-ite-FBg-8dwgiOXrJk4i5MJNo/edit	https://drive.goo
6	Tom Brian	44	Software Developer	95,000	https://docs.google.com/document/d/1HKX2DKw-MENG9K_wpzK0gTywcptRinnNf2XYJaXCuq0/edit	https://drive.goo
7	Winston Mckenzie	62	Artificial Intelligence Expert	150,000	https://docs.google.com/document/d/1sr3cW_xlztbaq1_BmpSME3VrAdbBEgFIBeK4ryzwEE/edit	https://drive.goo
8	Xiaoxi Chen	48	Senior Software Developer	115,000	https://docs.google.com/document/d/18qWW8dJulLm9bn6AZfnM4AWXyPmWVUY01xU2-bFncwl/edit	https://drive.goo
9	Albert Sandler	35	Research Assistant	80,000	https://drive.google.com/open?id=1EmawZotVNzD1TvP51t-wY1BlxIK5e9volb9Vy_Y6zE8	https://drive.goo
10	Alexadra Otero	41	Machine Learning Expert	134,000	https://drive.google.com/open?id=1dCrWbys5RQluXYCtRaZoq4iu0NXiIRXizAWt05KEQQU	https://drive.goo
11	Alicia Prochesko	44	Business Analyst	80,000	https://drive.google.com/open?id=1YthvGk3EY8sgGGZ_n2NqxVUlqUIB9PBM02l4WgInh5M	https://drive.goo
12	Alina Dawson	39	Research Assistant	70,000	https://drive.google.com/open?id=1uRBjbXgL9VQQV0uAALFMkBgluWrZ9aksFc8iu-jTRMg	https://drive.goo
13	Amalia Chater	54	Senior Software Developer	120,000	https://drive.google.com/open?id=1Xwx9rVjlQflhXtrGJ-HVZ83qxP4_WTTPO1wBipqXOl4	https://drive.goo

HR	/	EDIT LINKS
Doc	u	ments

Nes	Upload		Sync (	Share Nor		
All Dec	_	Find		Q		Adam Punjadi
- C	Name		Modified	Modified By	Tags	
~ ¢	Adah Punjabi		April 9, 201	8 🗆 Albin Ahme	ARRINET BACHELOR OF ENERGEBING BERLA INSTITUTE OF TECHNOLOGY C C++ MOTOROLA PROGRAM HANAGER PROGRAMMENG RESEARCH RESEARC	7360 GALLASSER DR APT 205 EDNA, MINNEAPOLIS, MN 55435 Yoshi: (952) 205-4598
q	Albert Sandler		April 9, 201	8 🗆 Albin Ahme	ARTIFICIAL INTELLISENCE COMMELL UNIVERSITY DATA STRUCTURE INITELLISENT SYSTEMANIKATY	E-mail: adub@projubi.edu
q	Alexandra Otero		April 9, 201	8 🗌 Albin Ahme	I DAGE PROCESSING ARROVE ARTIFICIAL INTELLIGENCE CAMBRIDGE CSS DATAM PROGRAMMING RESEARCH SOFTWARE ARCHITECT SUPPORT VECTOR MACHINE TO	ORACTIVE A full-time position in managing and designing business solutions to complex problems that requires entrepreneurial skills
¢	Alicia Prochesko		April 9, 201	8 🗌 Albin Ahme	C CLOUD COMPUTING CORPORATION NAVAGEMENT SALESFORCE SOFTWARE IN	coupled with solid technical knowledge and program management experience.
a	Alina Dawson		April 9, 201	8 🗆 Albin Ahme	CONSULTANT E-COMMERCE APPLICATION EDITOR EDUCATION IT MANAGEMENT	EDUCATION 2005-2007Stanford University, Computer Science Department, Stanford, CA
đ	Amalia Chater		April 9, 201	8 🗌 Albin Ahme	AMORE C C++ COMPUTER ARCHITECTURE DATABASE 3448, JAWASORPT 1	Master of Science, Expected graduation March 2007
q	Amber Smith		April 9, 201	8 🗆 Albin Ahme	A ADMINISTRATOR C DATABASE EDUCATION IT MANAGEMENT PROGRAMMING	1997-2001Birla Institute of Technology, Computer Science Department, Ranchi, India Bachelor of Engineering, Awarded Gold modal as the top ratiked candidate in the graduating class of 350
0	Arry Geller		April 9, 201	8 🗌 Albin Ahme	ANTIFICIAL INTELLISENCE CORPORATION EDUCATION INTELLISENT SYSTEMS LOW	PROFESSIONAL/RESEARCH EXPERIENCE
¢	Amy Wilson		April 9, 201	8 🗆 Albin Ahme	CONSULTION GOOGLE IT INVINGENENT MICROSOFT PROSERVING PRODE	01/2006-PresentResearch Assistant, Stanford Medical Informatics (SMI), Stanford University, Stanford, CA. Conducting research on a large database of HIV patients for finding temporal patterns among gene
q	Andreas Kahlenberg		April 9, 201	8 🗆 Albin Ahme	IT HWWGCHENT	mutations, drugs and elinical outcomes. • Developing and implementing multi-relational data mining algorithm for pattern discovery. This
6	Andy Hubs		April 9, 201	8 🗌 Albin Ahme	ANTERCON INTELLIGENCE C CONSIGNATION DEPOSITION ADDRESS OF DEPOSITION REP	will also support temporal querying, improved statistical aggregation and data visualization.

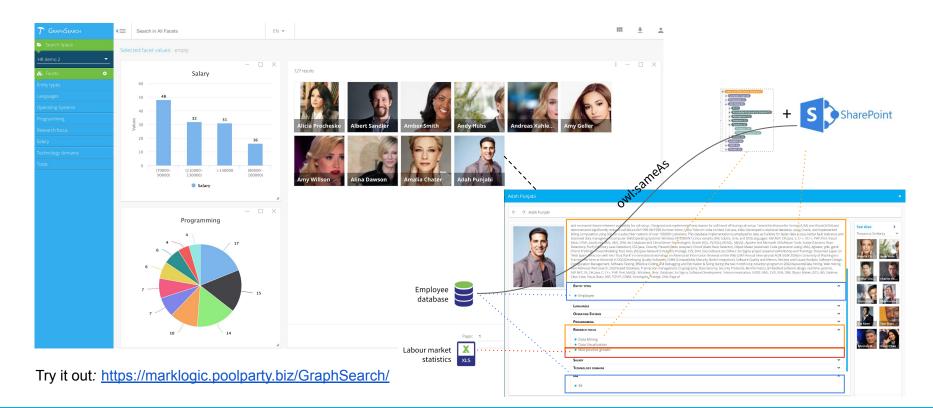
Programming					
Category	Skill	2015 Annual	2016 Annual ave	Growth Rate	
Programming	APL	509	505	-0.79%	
Programming	ASP.NET	9,288	9,398	1.18%	
Programming	С	130,629	135,629	3.83%	
Programming	C++	31,524	31,578	0.17%	
Programming	CSS	27,781	27,881	0.36%	
Programming	Eiffel	21	19	-9.52%	
Programming	HTML5	10,381	10,491	1.06%	

Category	Skil	2015 Annual avera	2016 Annual av	Growth Rate
Research focus	Data Mining	11,518	11,618	0.87%
Research focus	Data Visualization	13,294	13,394	0.75%
Research focus	Fuzzy Logic (FL)	71	69	-2.82%
Research focus	Genetic Algorithms (GAs)	1,783	1,785	0.11%
Research focus	Image Processing	11,856	11,866	0.08%
Research focus	Information retrieval	8,879	8,979	1.13%
Research focus	Intelligent agents	563	582	-0.18%
Research focus	Modeling Geometry	518	517	-0.19%
Research focus	Natural Language Processing	2,646	2,686	1.51%
Research focus	Neural Networks (NN)	1,110	1,120	0.90%
Research focus	Support Vector Machine (SVM)	270	271	0.37%
Research focus	Visual Datamining	5	5	0.00%



#### **Graph-based HR Analytics**

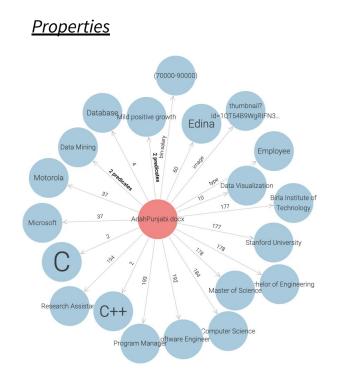




© Semantic Web Company

### Graph-based HR Analytics (cnt.)





<u>Note</u>: Predicates with numbers e.g. 2 represent the upper most concept (aka concept scheme) - "Programming", whereas "C" is the concept.

#### AdahPunjabi.docx &

X

AdahPunjabi.docx	
------------------	--

Types:

p120:Employee

RDF Rank:

0

Q Search instance properties

#### p120:age **39 <sup>xsd:long</sup>**

p120:gross-salary-per-year 80,000.00

dcterms:title Adah Punjabi

p120:jobTitle Research Assistant

p120:link-to-cv https://docs.google.com/document/d/1v9Q505



<u>Attributes</u>

© Semantic Web Company

- T

#### Graph-based HR Analytics (cnt.)



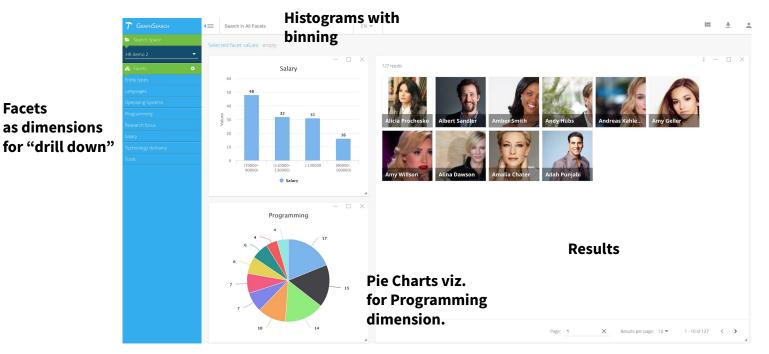
	subject 🗘	predicate 🜩	object 🗢
1	p0:Punjabi.docx	rdf.type	p120:Employee
2	p0:Punjabi.docx	dcterms:title	"Adah Punjabi"
3	p0:Punjabi.docx	p120:bin-salary	p120:70000
4	p0:Punjabi.docx	p120:gross-salary-per-year	"80,000.00"
5	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/10	p120:Skill01
6	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/10	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/129
7	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/10	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/159
8	p0:Punjabi.docx	p120:jobTitle	"Research Assistant"
9	p0:Punjabi.docx	p120:age	*39*^*xsd:long
10	p0:Punjabi.docx	p122:image	p123:QT54B9WgRIFN3a92CzpaB3uKyCVJIaTD
11	p0:Punjabi.docx	p120:link-to-cv	"https://docs.google.com/document/d/1v9Q505VC4rUm9VW897vX41U6ofD XVBrsqbIXhCjj80A/edit"
12	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/192	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/119
13	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/4	p120:Skill01
14	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/4	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/47
15	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/4	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/129
16	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/184	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/22
17	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/2	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/6
8	p0:Punjabi.docx	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/2	http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/7



#### **DESCRIBE** po:Punjabi.docx

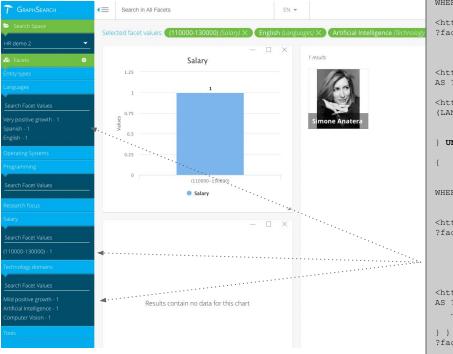
© Semantic Web Company

# Graph-based HR Analytics from the BI perspective

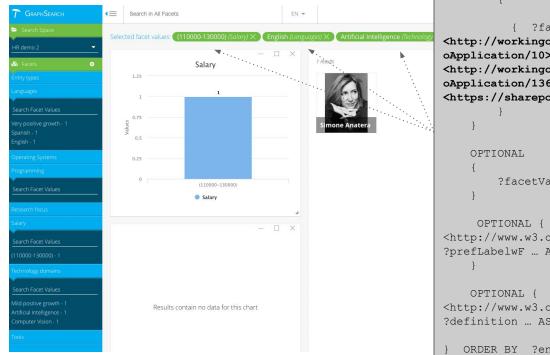


https://marklogic.poolparty.biz/GraphSearch/

poolparty<sub>®</sub>



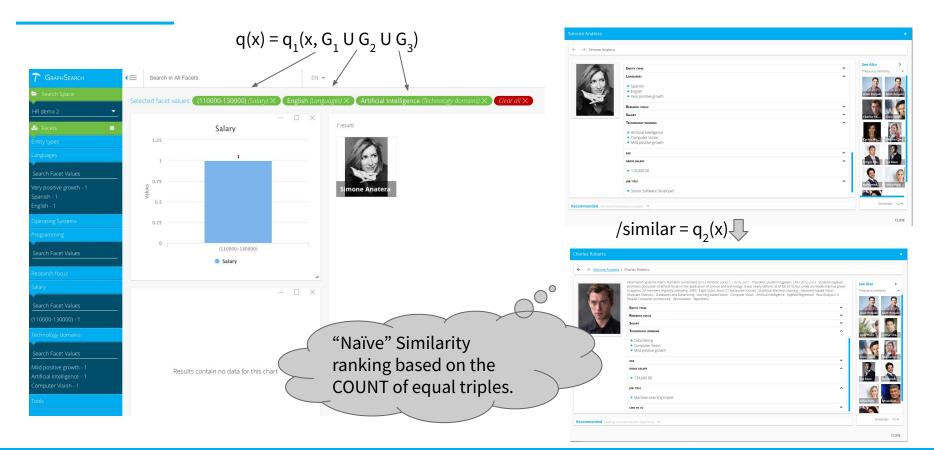
SELECT ?facetUri ?facetValueUri ?facetValueLabel ?facetCount FROM <http://hrdemo.com/example> WHERE { FILTER(?facetCount > 0).{ SELECT ?facetUri ?facetValueUri ?facetValueLabel ?facetCount WHERE SELECT (count(DISTINCT ?entity) as ?facetCount) ?facetValueUri WHERE { ?entitv <http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/4> ?facetValueUri . GROUP BY ?facetValueUri ORDER BY DESC(?facetCount) LIMIT 10 BIND ( <http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/4> AS ?facetUri ) . OPTIONAL { ?facetValueUri <http://www.w3.org/2004/02/skos/core#prefLabel> ?prefLabelIi.FILTER (LANG(?prefLabelIi) = '' || LANG(?prefLabelIi) = 'en') } UNION SELECT ?facetUri ?facetValueUri ?facetValueLabel ?facetCount WHERE SELECT (count(DISTINCT ?entity) as ?facetCount) ?facetValueUri WHERE ?entity <http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/3> ?facetValueUri GROUP BY ?facetValueUri ORDER BY DESC(?facetCount) LIMIT 10 BIND( <http://workingontologist.poolparty.biz/JobsandSkillsDemoApplication/3> AS ?facetUri ) . } } UNION { SELECT ?facetUri ?facetValueUri ?facetValueLabel ?facetCount WHERE



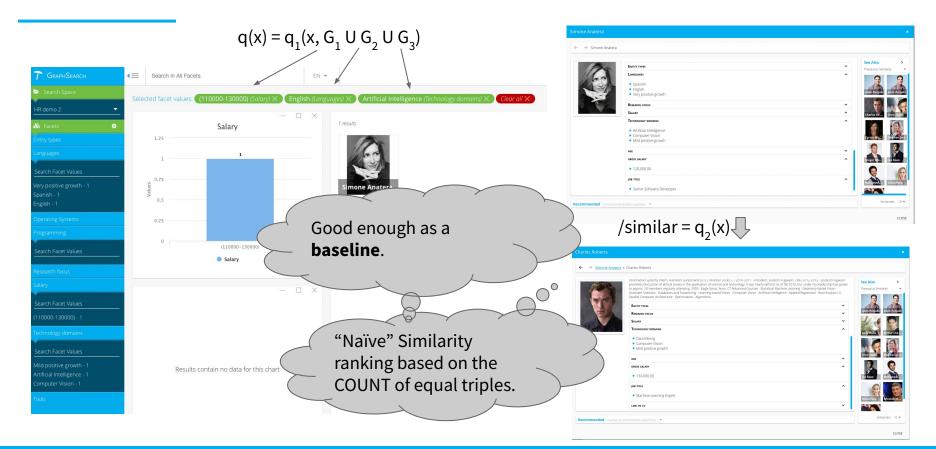
SELECT DISTINCT ?facetValueUri ?entityCreation ?entityTitle ?entityDescription FROM <http://hrdemo.com/example> WHERE SELECT DISTINCT ?facetValueUri WHERE ?facetValueUri <http://workingontologist.poolparty.biz/JobsandSkillsDem</pre> oApplication/10> ?object . } UNION { ?facetValueUri <http://workingontologist.poolparty.biz/JobsandSkillsDem</pre> oApplication/136> ?object . } UNION <https://sharepoint-demo/hr/bin-salary> ?object . } ?facetValueUri dct:created ?entityCreation ?facetValueUri <http://www.w3.org/2004/02/skos/core#prefLabel> ?prefLabelwF ... AS ?entityTitle) OPTIONAL { ?facetValueUri <http://www.w3.org/2004/02/skos/core#definition> ?definition ... AS ?entityDescription}

ORDER BY ?entityTitle LIMIT 10 OFFSET 0

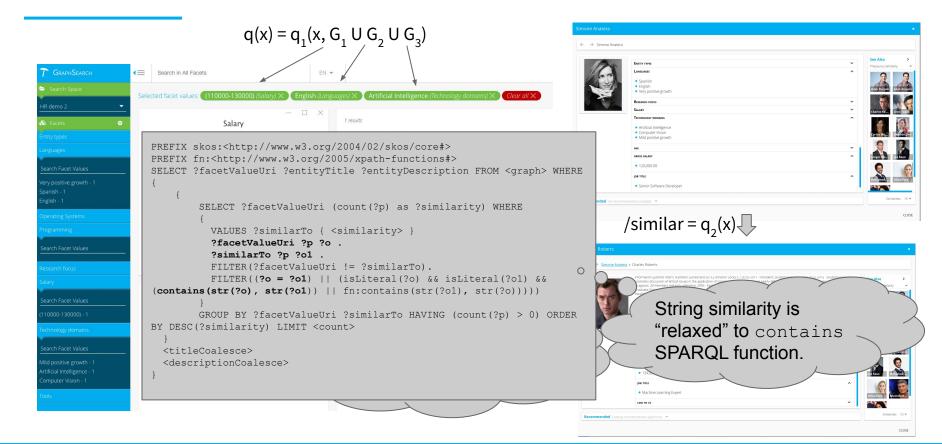






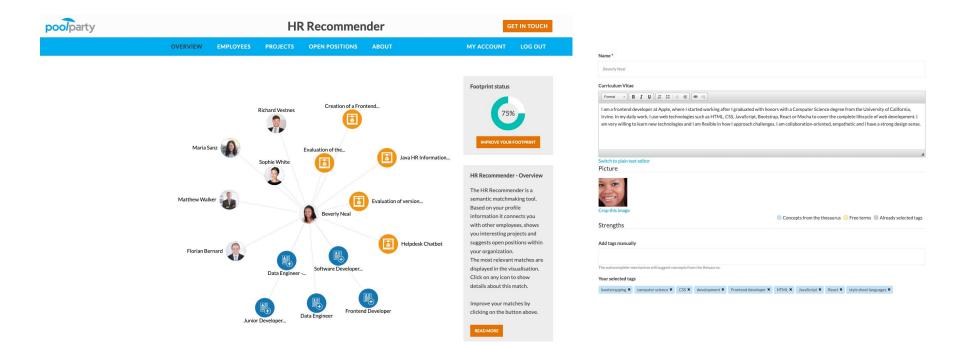






#### **KGs power Recommender Systems**

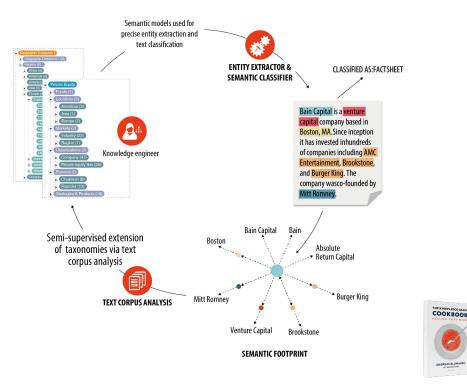




### Semantic footprint definition



**Definition**. Given an individual  $\mathbf{a}_1$  in ABox and a set of concepts  $\mathbf{c}$  in CBox that are returned from entity extraction, semantic footprint  $\oplus$  of  $\mathbf{a}_1$  encapsulates context of  $\mathbf{a}_1$  in respect to  $\mathbf{c}$  and CBox, i.e., concepts siblings, narrowers, broaders, and its custom relations, each with a respective weight  $\mathbf{w}$ .



#### KGs power Recommender Systems



		Give <b>weight w</b> based on a footprint. Use <b>boost b</b> for			
	HR Recommender	additional emphasis.		HR Recommender	
OVERVIEW EMPLOY	YEES PROJECTS OPEN POSITIONS ABOUT	and the second	OVERVIEW EM	APLOYEES PROJECTS OPEN POSITIONS ABOUT	
Meet these Emp Move the sliders to see the 	coworkers that best match your strengths - + computer science - + CSS + Frontend developer - + HTML - + React - + style sheet	RSET SLIDERS	Meet these I Move the sliders to see 	e the coworkers that best match your strengths ing - + computer science - + CSS	RESET SLIDERS
Employee Stre	ngths	q'(c) = q((c ⊕ w') * b	Employee	Strengths	
Matthew Java Walker	Script, HTML, C++, C#, PHP, web programming, Prolog, AJAX, Pascal, C	get in contact	Matthew Walker	JavaScript, HTML, C++, C#, PHP, web programming, Prolog, AJAX, Pascal, C	get in contact
Florian Bernard Java	Script, CSS, style sheet languages, Python, Jboss, ASP.NET, Prolog, AJAX,	get in contact	Florian Bernard	JavaScript, CSS, style sheet languages, Python, Jboss, ASP.NET, Prolog, AJAX, C	get in contact
Richard HTM Vestnes	AL, CSS, style sheet languages, Python, Jboss, ASP.NET, PHP, MATLAB	get in contact	Sophie White	JavaScript, CSS, style sheet languages, Java, C#, Scratch, JavaScript Framework, Perl	get in contact
	Script, CSS, style sheet languages, Java, C#, Scratch, JavaScript nework, Perl	get in contact	Richard Vestnes	HTML, CSS, style sheet languages, Python, Jboss, ASP.NET, PHP, MATLAB	get in contact
Maria Sanz From	ntend developer, Python, C#, PHP, Delphi, AJAX, Perl, Objective-C	get in contact	Maria Sanz	Frontend developer, Python, C#, PHP, Delphi, AJAX, Perl, Objective-C	get in contact
	, style sheet languages, Python, PHP, MATLAB, web programming, ective-C	get in contact	Erico Ramos	<b>computer science, JavaScript,</b> PHP, JavaScript Framework, Pascal, Perl, Objective-C	get in contact
	puter science, JavaScript, PHP, JavaScript Framework, Pascal, Perl, active-C	get in contact	Lewis Wright	CSS, style sheet languages, Python, PHP, MATLAB, web programming, Objective-C	get in contact
	Script, Python, ASP.NET, JavaScript Framework, AJAX, Objective-C, grated development environment software	get in contact	Donna Moreno	HTML, Java, Python, Prolog, JavaScript Framework, AJAX, C	get in contact
Donna Moreno HTM	ML, Java, Python, Prolog, JavaScript Framework, AJAX, C	get in	Nathaniel	JavaScript, Python, ASP.NET, JavaScript Framework, AJAX, Objective-C	get in

### KGs power Recommender Systems (cnt.)



oolparty		HR Recommender	GET IN TOUCH
	OVERVIEW EMPLOYEES PI	OJECTS OPEN POSITIONS ABOUT	MY ACCOUNT LOG OUT
	Interesting Projects for Move the sliders to see the projects that - + bootstrapping - - + development - - + JavaScript -		RESET SLIDERS Footprint status
	Project name	Keywords	IMPROVE YOUR FOOTPRINT
	Java HR Information Extractor	React, Frontend developer, Natural language processing, Java, Java Software Developer	get in contact
	Creation of a Frontend Style Guide	development, CSS, style sheet languages, Frontend developer	get in contact Don't miss out on interesting
	E Helpdesk Chatbot	React, Java, Python	get in projects anymore! Here you find projects that match your strenghts and interests.The
	Evaluation of the current NLP status	Natural language processing, Java, Python, Haskell, Lisp	get in contact         recommendation is based on your footprint and may be fine-
	Evaluation of version control tools	development, computer science, computer programming	get in veaken or boost a strength. contact Direct matches are displayed in
	Workshop: Data structures and algorithms	development, computer science	get in         black and matches derived from           contact         the knowledge graph in gray.
	A Framework for Publishing and Consuming Data Asse	Java, Python, C++	get in Improve your matches by contact

Difference of **recommendation** with respect to **similarity**: Recommender systems not necessarily recommend within the same entity type, e.g. jump from Employees to Projects.

#### **Boosted + semantic footprint scoring**



```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
PREFIX esco: <http://data.europa.eu/esco/model#>
SELECT ?uri ?finalScore
WHERE {
 BIND(?uri AS ?uri) .
   SELECT ?uri ?finalScore
   WHERE
     {
        SELECT ?uri (MAX(?boostedScore) as ?maxDistScore) (MAX(?selfDistScore) as ?selfScore)
        WHERE {
          VALUES (?x ?boost) { <inputConcepts> }
            BIND(?x AS ?uri)
               BIND(STRDT('1.00', xsd:float) AS ?distScore)
         ) UNION (
               ?x skos:narrower ?y.
               BIND(STRDT('0.8',xsd:float) AS ?distScore)
           UNION {
               ?x <http://vocabulary.semantic-web.at/cocktail-ontology/consists-of> ?y .
               BIND(STRDT('0.7', xsd:float) AS ?distScore)
           }
         BIND(IF(BOUND(?y), ?y, ?x) AS ?uri) .
         BIND(?distScore * ?boost * 0.5 AS ?boostedScore) .
         BIND(IF(BOUND(?v), ?empty, ?distScore * ?boost) AS ?selfDistScore) .
        1
        GROUP BY ?uri
      BIND(IF(BOUND(?selfScore), ?selfScore, ?maxDistScore) AS ?finalScore) .
ORDER BY DESC(?finalScore)
LIMIT <numberOfConcepts>
```

```
Semantic footprint ⊕:
Actual concept: 1
narrowers: 0.8
consists-of: 0.7
```

*Boost* is the input variable from the GUI.

#### **Demo: Semantic matchmaking**



#### Consultants with best match for selected project

#### **Project title**

Elucidating the molecular mechanisms of age-related muscle loss to inform Artificial Intelligence-based discovery of novel peptide therapeutics for sarcopenia.

#### Semantic footprint of project

accelerator, Accounting, age, age-related, ageing, Artificial Intelligence, clinical trials, complexity, data, develop, disease, disorders, driving, Human, industry, injury, knowledge, modelling, models, muscle, peptide, Researcher, systems, therapies

#### Enhance matching with

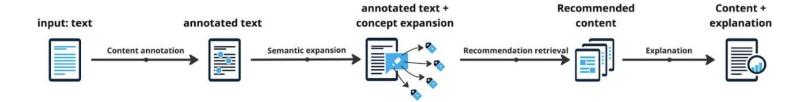
Enrich footprint using the Knowledge Graph	
Consider experience level based on structured data	2

Consultant	Strengths	Score
Joel Mccoy	data, develop, Automation Specialist, metadata	100
Denil Pereira	data, develop, Architect, Biology	97
Christopher Mcdonalid	data, systems	57
Eddie Ferguson	data, Human	57
Jackson White	Artificial Intelligence, driving	57

#### Step by step demo: <u>https://semantic-matchmaker.poolparty.biz/</u>

#### **Recommender workflow**





### Semantic Search/Recommenders vs Zettelkasten



- Zettelkasten technique can be implemented in GraphSearch:
  - Categorized search (using keywords #) is done via dereferenceable concepts, e.g., <a href="http://dbpedia.org/page/Leonardo\_da\_Vinci">http://dbpedia.org/page/Leonardo\_da\_Vinci</a>, <a href="http://dbpedia.org/page/Category:Renaissance">http://dbpedia.org/page/Category:Renaissance</a>>
  - 2. Regarding **related index cards** instead of pointers [[]] we have **typed links**, dereferenceable properties, e.g., <<u>http://dbpedia.org/ontology/museum</u>>
  - 3. For related index cards, we can use **similarity or recommend** feature based on *semantic footprint*.

### Dereferenceable entities (Things, not strings)



• Entities (concepts or properties) are dereferenceable so that one can "look up" the entity, understand the context and its meaning.

dbpedia.org/page/L-	🔒 dbpedia.org/page/Leonardo_da_Vinci 🔍 🖒 🏠		dbpedia.org/ontology/museum		G		
	💐 DBpedia 🔹 Browse using * 🗈 Formats * 🔅 Faceted Browser 🕑 Sparqi Endpoint		= aspealatorg/entere				
		nardo da Vinci , from Named Graph. <u>http://diceedia.org</u> , within Data Space: <u>gbreedia.org</u>			🔆 DBpedia 🔹 🕫	rowse using 👻 📓 Formats 👻	🖉 Faceted Browser 🛛 Sparql Endpoint
Leonardo di ser Piero da Vinc as a painter, draughtsman, er achievements as a painter, hu including anatorny, astronom		o da Vinci (15 April 1452 – 2 May 1519) was an Italian polymath of the High Renatisance who was active same, verginera, scientific, theoriti, acutoric, and activitac. While his fane hitality rested on his antice, he airo bacema from koris, hin which he made dawing and refores on a vertedy of a Jubicst, attornomy, buttom, cantorgrarky, parting, and patienticlogy, Lenando is widdly regarded to have been a ref de Pensiuscence humania false, and its Gorcicke works comprise activitation tables.			About: <u>museum</u> An Entity of Type: <u>Property</u> , from Named Graph: <u>http://dbpedia.org/resource/classes#</u> , within Data Space: <u>dbpedia.org</u>		
	of artists matched only by that of his younger contemporary, Michelangelo.		and the second se		Property	Value	
		a compared and a			rdf: <u>type</u>	rdf <u>-Property</u> oxt:ObjectProperty	
	Property	Value					
	stor <u>absitast</u>	• Lowards die Price Arb (c): 15 April 142; - 2 May (2)) saus in balan polyvatent für die High Therasakance Arb was achte aus a participation of the second of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus a participation of the High Therasakance Arb was achte aus achter Arb Was and Was and Was and Arb Was and Was and Was and Arb Was and			rdfs:domain	dbo:Artwork	
					rdfs:isDefinedBy	<ul> <li>http://dbpedia.org/ontology/</li> </ul>	
					rdfs;label	museum (en)	
					rdfs:range	dbo;Museum	
					rdfs:subPropertyOf	dui <u>:hasLocation</u>	
					wdrs:describedby	dbe <u>:data/definitions.ttl</u>	
					prov:wasDerivedFrom	http://mappings.dbpedia.org/index.php/OntologyProperty:museum	
	dbo.birthOate	<ul> <li>1452-04-15 (xsdxlate)</li> </ul>			is ov:defines of	http://dbpedia.org/ontology/	
	doxbirthName	Leonardo di ser Piero da Vinci (en)					
	dbo.birthPlace	dic <u>Republic of Florence</u> dickechilop			is ov:describes of	dbo <u>:data/definitions.ttl</u>	

### Semantic Search/Recommenders vs Zettelkasten



"From the perspective of search — in this system, one can either "**branch**" further via categories, or "**chase**" using pointers to related index cards."

"By combining "branch" and "chase" one can find information that on first look appeared disconnected, but turned out to be very relevant information."

	Information Systems Intern, Hamilton Sundstrand (UFC), Windsor Locks, CT 2010-2011 - President, Student Pugwash, CMU 2012-2013 - Student Pugwash	
(	protects discussion of chical issues in the application of science and technology it was nearly different as of fail 2010, but under my leadership has grown in agence. 30 members regularly among 2003 Faight Science, Novel CT Advanced Consers: Statistical Matchine Learning, Genemetry-band Vision Graduate Satistics: Databases and Datamining. Learning based Vision: Computer Vision. Artificial intelligence: Applied Regression: Real Analysis 14- Paralel (Computer Archinecture Ogmitation: Agenthem)	See Also Thesaurus Similarity
	ENTITY TYPES Y	dia d
	Research rocus v	Adah Punjabi Adah Pu
	Salaar V	4 6 5-
	TECHNOLOSY DOMAINS	
	Duta Mining     Computer Vision     Mid postive growth	
	462 V	Chen Yuang Elvavor
	ERGS 564APT ^	
	je mi	
	Machine Learning Expert	Olivia Park
	LINK TO CV V	



OVEMBER 27, 2019

← Back to blog

There is a strong connection between information storage and retrieval, namely how information storage is implemented under the hood directly impacts the retrieval or knowledge search.

The knowledge management techniques we adopt have an immense impact on our daily productivity. For that reason, a lot of thought is given by us when we store information, asking questions such as which tools we should use for the very same purpose.

In this blog post, we will see the reasons behind why PoolParty GraphSearch is a very powerful semantic and knowledge search tool, but before dwelling into that we go back to an old-school method of knowledge management called "Zettellasten".



Albin Ahmeti DATA KNOWLEDGE ENGINEE in

## Semantic Search/Recommenders vs Zettelkasten



- Zettelkasten technique can be implemented in GraphSearch:
  - 1. **Categorized search** (using keywords #) is done via dereferenceable concepts, e.g., <a href="http://dbpedia.org/page/Leonardo\_da\_Vinci">http://dbpedia.org/page/Category:Renaissance</a>>
  - 2. Regarding **related index cards** instead of pointers [[]] we have **typed links**, dereferenceable properties, e.g., <<u>http://dbpedia.org/ontology/museum</u>>
  - 3. For related index cards, we can use **similarity or recommend** feature based on *semantic footprint*.

**Proposition.** In (PoolParty) GraphSearch 2) and 3) *extend* the Zettelkasten technique, while 1) makes the approach, in general, more robust, i.e., representing data using "things" and not "strings".

## **Closing remarks: URIs management**



- The following are assumptions underlying the URIs:
  - URIs should not change;
  - If the source authority provides URIs (file is RDF) we will preserve those, and by clicking those URIs it will forward you to the authority frontends to get more details on them - e.g. <u>http://eunis.eea.europa.eu/habitats/409</u>;
  - If the source authority provides no URIs (file is non-RDF, such as CSV), we will create URIs that adhere to the project URI generation mechanism, that will make those URIs de-referenceable on PoolParty frontend e.g.

https://sensingclues.poolparty.biz/HabitatClassificationScheme/237

## **URIs** management

ensingclues.poolparty.biz/HabitatClassificationScheme/237.html	Q	Û	☆		•
				SPARQL	. 13
A B C D E F G H I J K L M N O P Q R S T	U V	/ W	X	Y	Z
PoolParty Linked Data Server - Taxonomies, Thesauri, Vocabularies > IUCN Habitat Classification Schen	ne			EN	*
HTML VISUAL					
Artificial - Aquatic					
https://sensingclues.poolparty.biz/HabitatClassificationScheme/237					
Notation					
15 Artificial - Aquatic					
Narrower Concepts					
Aquaculture Ponds					
Canals and Drainage Channels, Ditches					
Excavations (open)					
Irrigated Land [includes irrigation channels]					
Karst and Other Subterranean Hydrological Systems [human-made]					
Mari/Brackish-culture Ponds					
Mariculture Cages					
Marine Anthropogenic Structures					
Ponds [below 8 ha]					
Salt Exploitation Sites					
Seasonally Flooded Agricultural Land					
Wastewater Treatment Areas					
Water Storage Areas [over 8 ha]					

	$\leftarrow \   \rightarrow$	C 🔒 eunis.eea.europa.eu/habitats/409								
		JNIS Home Species Habitat types Sites Global queries References About EUNIS								
1		UNIS habitat classification 2012 amended 2019 > B - Coastal habitats								
		oastal habitats								
		English name: Coastal habitats Quick facts								
		Description (English)	EUNIS habitat type code B							
mptions ur		Coastal habitats are those above spring high tide limit (or above mean water level in non-tidal waters) occupying coastal features and characterised by their								
		proximity to the sea, including coastal dunes and wooded coastal dunes, beaches and cliffs. Includes free-draining supralittoral habitats adjacent to marine habitats								
2,		which are normally only affected by spray or splash, strandlines characterised by terrestrial invertebrates and moist and wet coastal dune slacks and dune-slack								
		pools. Excludes supralittoral rock pools and habitats adjacent to the sea which are not characterised by salt spray, wave or sea-ice erosion.								
aravidas UDV		Source: EUNIS habitat classification								
provides URIs	5 (11	te is RDF) we will preserve	those, and							
t will forwarc	l yo	ou to the authority frontend	ls to get							
e.g. <u>http://eunis.</u>	eea	europa.eu/habitats/409;								

provides no URIs (file is non-RDF, such as CSV), we will to the project URI generation mechanism, that will

- make those URIs de-referenceable on PoolParty frontend e.g.
- https://sensingclues.poolparty.biz/HabitatClassificationScheme/237

R

## URIs management (cnt.)



sensingclues.poolparty.biz/HabitatClassificationScheme/237.html	Q 🖞 🖞		*	$\leftarrow$ $\rightarrow$ C $\bullet$ sensingclues.poolparty.biz/HabitatClassificationScheme/237.ttl
A B C D E F G H I J K L M N O P Q R S T	UVWX	Y Z		<pre>@prefix rdf: <http: 02="" 1999="" 22-rdf-syntax-ns#="" www.w3.org=""> . @prefix rdfs: <http: 01="" 2000="" rdf-schema#="" www.w3.org=""> .</http:></http:></pre>
PoolParty Linked Data Server - Taxonomies, Thesauri, Vocabularies > IUCN Habitat Classification Scheme		EN ¥		<pre>@prefix skos: <http: 02="" 2004="" core#="" skos="" www.w3.org=""> . @prefix skoss! <http: 05="" 2008="" skos-x1#="" www.w3.org=""> . @prefix skoss! <http: 05="" 2008="" skos-x1#="" www.w3.org=""> . @prefix ow1: <http: 05="" 2008="" skos-x1#="" www.w3.org=""> .</http:></http:></http:></http:></pre>
HTML VISUAL				<pre>%prefix dc: <http: 1.1="" dc="" elements="" purl.org=""></http:> . %prefix dcterms: <http: dc="" purl.org="" terms=""></http:> .</pre>
Artificial - Aquatic				<pre>@prefix csw: <http: csw.owl#="" ontologies="" semantic-web.at=""> . @prefix ctag: <http: commontag.org="" ns#=""> . @prefix opencyc: <http: concept="" sw.opencyc.org=""></http:> .</http:></http:></pre>
https://sensingclues.poolparty.biz/HabitatClassificationScheme/237				<pre>@prefix freebase: <htp: 2001="" www.w3.org="" xmlschema#=""> .</htp:></pre>
Notation				<pre>%prefix tags: <http: 0.1="" owl="" redwood="" tags="" www.holygoat.co.uk=""></http:> . %prefix foaf: <http: 0.1="" foaf="" xmlns.com=""></http:> .</pre>
15 Artificial - Aquatic				<pre>@prefix cyc: <http: coconcept="" sw.cyc.com=""></http:> . @prefix cycannot: <http: cycannotations_vl#="" sw.cyc.com=""> . @prefix dbpedia: <http: dbpedia.org="" resource=""></http:> .</http:></pre>
Narrower Concepts				<pre></pre> <https: 237="" habitatclassificationscheme="" sensingclues.poolparty.biz=""> a skos:Concept;  dcterms:created "2022-11-24T14;42:49.3402"*xad:dateTime;</https:>
Aquaculture Ponds				<pre>skos:prefLabel "Artificial - Aquatic"@en; skos:topConceptOf <https: 127="" habitatclassificationscheme="" sensingclues.poolparty.biz="">;</https:></pre>
Canals and Drainage Channels, Ditches				<pre>skos:notation "15 Artificial - Aquatic"; skos:narrower <https: 238="" habitatclassificationscheme="" sensingclues.poolparty.biz="">,</https:></pre>
Excavations (open)				<https: 239="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 240="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 241="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 242="" habitatclassificationscheme="" sensingclues.poolparty.biz="">,</https:></https:></https:></https:>
Irrigated Land [includes irrigation channels]				<pre><https: 243="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 245="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 245="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 246="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 247="" habitatclassificationscheme="" sensingclues.poolparty.biz="">, <https: 248="" habitatclassificationscheme="" sensingclues.poolparty.biz="">,</https:></https:></https:></https:></https:></https:></pre>
Karst and Other Subterranean Hydrological Systems [human-made]				<pre></pre>
Mari/Brackish-culture Ponds				
Mariculture Cages				
Marine Anthropogenic Structures				
Ponds [below 8 ha]				
Salt Exploitation Sites				Asservations to the Linds of Dester Deins sinds a need Time Desmands
Seasonally Flooded Agricultural Land				According to the Linked Data Principles and Tim Berner Lee
Wastewater Treatment Areas				Semantic Web vision, to agents is served RDF, whereas to
Water Storage Areas [over 8 ha]				
			<u> </u>	people human-readable data.

https://sensingclues.poolparty.biz/HabitatClassificationScheme/237

## URIs management (cnt.)

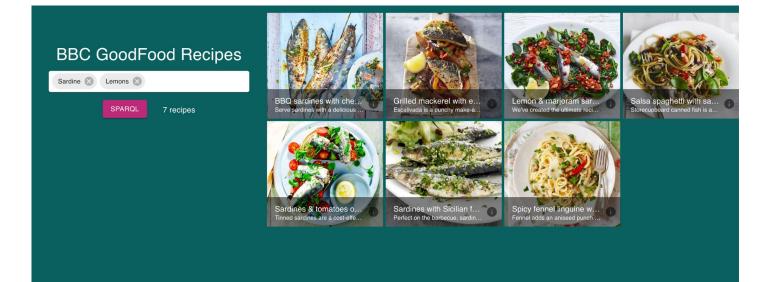




- 1 curl --location 'https://sensingclues.poolparty.biz/ HabitatClassificationScheme/237' \
- 2 --header 'Accept: application/rdf+xml'

## Part II: The recipe











• We are going to use the following suite of tools..*though the principles discussed here can be generalized.* 



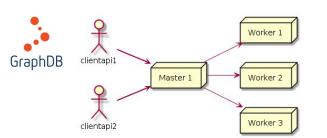
RDFOX

(UnifiedViews, Extractor, Thesaurus Manager)



## PP Architecture and scaling with big data

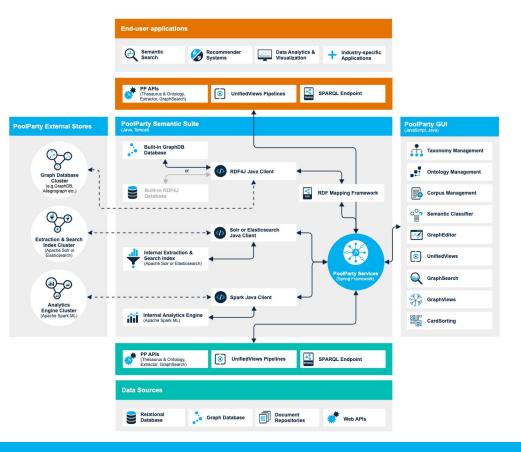




(persistent, high-availability, scale-out)



(in-memory\*, high-availability)



## Integrating taxonomies for information retrieval



- A certain number of taxonomies are already published as RDF, or they are accessible via SPARQL endpoints.
  - Only a subset of them are SKOS-compliant.
  - We have to resort to SPARQL update queries, or use tools like PoolParty Data Validator to repair them.
- Otherwise they are still stored in XLS or CSV format, so we have to convert them to a format which is SKOS-compliant.
  - We have to map XLS/CSV to RDF via a set of mappings.

### .. unless our tool supports ingesting XLS/CSV directly!

## **Google Product Taxonomy integration** [5]



• **Google Product Taxonomy (GPT)** is a taxonomy used by Google in categorizing products, for the purpose of ensuring that a certain advertisement is <u>recommended</u> with the search results.

# Google_Product_Taxonomy_Version: 2021-09-21
Animals & Pet Supplies
Animals & Pet Supplies > Live Animals
Animals & Pet Supplies > Pet Supplies
Animals & Pet Supplies > Pet Supplies > Bird Supplies
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Cage Accessories
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Cage Accessories > Bird Cage Bird Baths
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Cage Accessories > Bird Cage Food & Water Dishes
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Cages & Stands
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Food
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Gyms & Playstands
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Ladders & Perches
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Toys
Animals & Pet Supplies > Pet Supplies > Bird Supplies > Bird Treats
Animals & Pet Supplies > Pet Supplies > Cat Supplies
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Apparel
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Beds
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Food
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Food > Non-prescription Cat Food
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Food > Prescription Cat Food
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Furniture
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Furniture Accessories
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Litter
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Litter Box Liners
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Litter Box Mats
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Litter Boxes
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Toys
Animals & Pet Supplies > Pet Supplies > Cat Supplies > Cat Treats
Animals & Pet Supplies > Pet Supplies > Dog Supplies
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Apparel
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Beds
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Diaper Pads & Liners
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Diapers
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Food
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Food > Non-prescription Dog Food
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Food > Prescription Dog Food
Animals 4 Pet Supplies > Pet Supplies > Dog Supplies > Dog Houses
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Kennel & Run Accessories
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Kennels & Runs
Animals & Pet Supplies > Pet Supplies > Dog Supplies > Dog Toys
Animals 4 Pet Supplies > Pet Supplies > Dog Supplies > Dog Treadmills
Animals 4 Pet Supplies > Pet Supplies > Dog Supplies > Dog Treats
Animals & Pet Supplies > Pet Supplies > Fish Supplies
Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aguarium & Pond Tubing
Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium & Pond Tubing Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium Air Stones & Diffusers
Animals & Fet Supplies > Fet Supplies > Fish Supplies > Aquarium All Scholes & Diffuses Animals & Fet Supplies > Fet Supplies > Fish Supplies > Aquarium Cleaning Supplies
Animals & Fet Supplies > Fet Supplies > Fish Supplies > Aquarium Cleaning Supplies
Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium Decor
Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium Filters
Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium Fish Mets Animals & Pet Supplies > Pet Supplies > Fish Supplies > Aquarium Fish Mets
Anishis & Fet Substres / Fet Substres / Addition Graver & Substrates

en

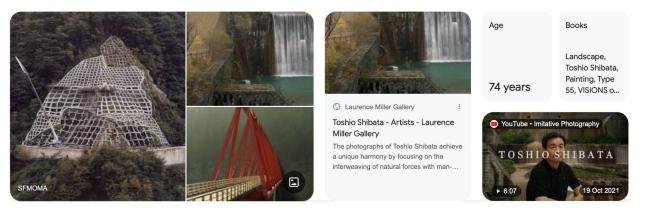
# Google\_Product\_Taxonomy\_Version: 2021-09-21 Baby & Kleinkind Baby & Kleinkind > Baby Badebedarf Baby & Kleinkind > Baby Badebedarf > Babybadewannen Baby & Kleinkind > Baby Badebedarf > Shampoo-Schutzschilder Baby & Kleinkind > Baby Gesundheitsbedarf Baby & Kleinkind > Baby Gesundheitsbedarf > Baby-Gesundheits- & Pflegesets Baby & Kleinkind > Baby Gesundheitsbedarf > Nasensauger Baby & Kleinkind > Baby Gesundheitsbedarf > Reinigungstähcher fähr Schnuller Baby & Kleinkind > Baby Gesundheitsbedarf > Schnuller & BeiÄŸringe Baby & Kleinkind > Baby Gesundheitsbedarf > Schnullerclips & -ketten Baby & Kleinkind > Baby Gestmanere Baby & Kleinkind > Babysicherheit > Baby- & Haustierschutzgitter Baby & Kleinkind > Babysicherheit > Baby-Schutzleisten Baby & Kleinkind > Babysicherheit > Babyfone Baby & Kleinkind > Babysicherheit > Babysicherheitsgurte & -leinen Baby & Kleinkind > Babysicherheit > Babysicherungen & Schutzvorrichtungen Baby & Kleinkind > Babysicherheit > ZubehÄfr fÄhr Baby- & Haustierschutzgitter Baby & Kleinkind > Babyspielwaren Baby & Kleinkind > Babyspielwaren > ABC-Lernspielzeuge Baby & Kleinkind > Babyspielwaren > Baby-Aktiv-Spielzeug Baby & Kleinkind > Babyspielwaren > Baby-Mobile-ZubehÄir Baby & Kleinkind > Babyspielwaren > Baby-Mobiles Baby & Kleinkind > Babyspielwaren > Babylauflernhilfen Baby & Kleinkind > Babyspielwaren > Babyrasseln Baby & Kleinkind > Babyspielwaren > Babyschaukeln & TÄhrhopser Baby & Kleinkind > Babyspielwaren > Babywippen Baby & Kleinkind > Babyspielwaren > Lauf- & Spielgitter Baby & Kleinkind > Babyspielwaren > Schnuller & Beruhigung Baby & Kleinkind > Babyspielwaren > Sortier-, Stapel- & Steckspielzeug Baby & Kleinkind > Babyspielwaren > Spiel- & Krabbeldecken Baby & Kleinkind > Babyspielwaren > Spiel- & Krabbeldecken > Krabbeldecken Baby & Kleinkind > Babyspielwaren > Spiel- & Krabbeldecken > Spieldecken Baby & Kleinkind > Babyspielwaren > Zieh- & Schiebespielzeug Baby & Kleinkind > Babytransport Baby & Kleinkind > Babytransport > Baby- & Kleinkindautositze Baby & Kleinkind > Babytransport > Babytransport > Baby & Kleinkind > Babytransport > Kinderwager Baby & Kleinkind > BabytransportzubehÄ%r Baby & Kleinkind > Babytransportzubehā%r > Einkaufswagen- & Hochstuhlbezākge Baby & Kleinkind > BabytransportzubehĶr > FuÄŸsĤcke Baby & Kleinkind > Babytransportzubehā%r > Zubehā%r fā%r Baby- & Kleinkindautositze Baby & Kleinkind > Babytransportzubehā%r > Zubehā%r fā%r Babytrā#ger Baby & Kleinkind > BabytransportzubehÄ%r > ZubehÄ%r fÄ%r Kinderwagen Baby & Kleinkind > Geschenksets fåbr Babys Baby & Kleinkind > Puckdecken Baby & Kleinkind > Puckdecken > Einschlagdecken Baby & Kleinkind > Puckdecken > Einschlagdecken & Puckdecken Baby r Floinkind > Stillon r Pälttown

## **Google Product Taxonomy**



Toshio Shibata

Japanese photographer



Search for the photographer "Toshio Shibata"

0 ()

Laurence Miller Gallery
 http://www.laurencemillergallery.com > artists > toshio...

#### Toshio Shibata - Artists

The photographs of **Toshio Shibata** achieve a unique harmony by focusing on the interweaving of natural forces with man-made structures.

#### shashasha

https://www.shashasha.co > artist > toshio-shibata

#### Toshio SHIBATA - 柴田敏雄 | shashasha 写々者

Born in 1949 in Tokyo, **Toshio Shibata** has been producing a body of photographs which both question and recreate the image as landscape.





#### About

Toshio Shibata is a Japanese photographer known for his large-format photographs of large-scale works of civil engineering in unpopulated landscapes. Shibata was born in Tokyo. He graduated from Tokyo University of the Arts with a B.A. in 1972 and an M.F.A. in 1974 in which he concentrated primarily on painting. Wikipedia

Born: 1949 (age 74 years), Tokyo, Japan

**Books:** Landscape, Toshio Shibata, Painting, Type 55, MORE

Results retrieved from the *Google Knowledge Graph*.

## **Google Product Taxonomy**



#### Toshio Shibata

Japanese photographer



Laurence Miller Gallery
 http://www.laurencemillergallery.com > artists > toshio...

#### Toshio Shibata - Artists

The photographs of **Toshio Shibata** achieve a unique harmony by focusing on the interweaving of natural forces with man-made structures.

#### shashasha

https://www.shashasha.co > artist > toshio-shibata

#### Toshio SHIBATA - 柴田敏雄 | shashasha 写々者

Born in 1949 in Tokyo, **Toshio Shibata** has been producing a body of photographs which both question and recreate the image as landscape.

#### Gale Toshio Shibata

Toshio Shibata is one of Japan's most respected contemporary photographers. His work has been the subject of exhibitions and publications throughout Japan, the ... £90.00  $\cdot$  In stock

#### Phillips https://www.phillips.com / artist / toship-shibata

#### Toshio Shibata - Phillips

Browse the works and learn more about **Toshio Shibata**. Find upcoming and past auctions and exhibitions featuring their work at Phillips.

#### Sponsored

eBay
 https://www.ebay.com

#### Toshio shibata books for sale | eBay

Free Shipping Available. Buy Top Products on eBay, Money Back Guaranteel Buy It Now Available. Huge Selection. Buyer Protection Program. Featured Collections. Daily Deals. Under \$10 - Sell on eBay - Trending on eBay - eBay Money Back Guarantee - Daily Deals.

#### Sponsored

Fuji Arts http://www.fujiarts.com

Fuji Arts Japanese Prints - Affordable Worldwide Shipping Shop Online Now- New auctions daily! The highest volume Japanese print retailer. Fuji Arts Japanese Woodblock Prints. Great service, terrific prices. Shop now! Over 2,500 prints. Fine Woodblock Reprints - Edo Era, 1603 - 1868 - Clearance Bargains - Just Listed Yoshtoshi prints - from US\$98.00 - Meiji period master - More \*

#### Sponsored

S mutualart.com

https://www.mutualart.com

#### Toshio Shibata | 41 Exhibitions and Events | MutualArt

Receive alerts about your favorite trending artists and art venues to never miss a sale. Stay up to date with art market news, trends, events, emerging artists and sale outcomes. Art Info. Online Art Appraisal - Exhaustive Price Database - Art Auctions - Sell Your Art - MyCollection Summer Sale: 20% off PREMUM Annual Plan

Some results may have been removed under data protection law in Europe. Learn more

#### 

Search for the photographer "Toshio Shibata"



## **Google Product Taxonomy schema**



• PoolParty allows to import XLS/CSV, but one has to adhere to a **target schema**.

A	В	С	D	E	F	G
5	Animals & Pet Supplies					
3237	Animals & Pet Supplies	Live Animals				
2	Animals & Pet Supplies	Pet Supplies				
5	Animals & Pet Supplies	Pet Supplies	Bird Supplies			
7385	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories		
499954	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories	Bird Cage Bird Bat	hs
7386	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories	Bird Cage Food &	
4989	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cages & Stands	and dage i dod a	
4990	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Food		
7398	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Gyms & Playstands		
4991	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Ladders & Perches		
4992	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Toys		
4993	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Treats		
4	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Bid Houts		
5082	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Apparel		
4433	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Beds		
3367	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food		
543684	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food	Non-prescription C	at Food
543683	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food	Prescription Cat Fo	
4997	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Fumiture	r tooonpaon out re	
500059	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Furniture Accessories		
4999	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter		
8069	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Box Liners		
7142	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Box Mats		
5000	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Boxes		
5001	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Toys		
5002	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Treats		
5	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Cat fields		
5004	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Apparel		
4434	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Beds		
7372	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Diaper Pads & Liners		
499900	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Diapers		
3530	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food		
543682	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food	Non-prescription D	ng Eood
543681	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food	Prescription Dog F	
5094	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Houses	r reacription bog P	
7428	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Kennel & Run Accessories		
7274	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Kennels & Runs		
5010	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Toys		
8123	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Treadmills		
5011	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Treats		
5011	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Sog freats		
505303	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium & Pond Tubing		
505303	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Air Stones & Diffuser		
500038	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Air Stones & Dirusers Aquarium Cleaning Supplies	,	
500038	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Decor		
5019	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Electr		
5020	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Fish Nets		
505306	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Fish Nets Aquarium Gravel & Substrates		

#### source schema

#### target schema

Animals & Pet Supplies					
P	et Supplies				3
		Pet Sunscreen			7396
		Vehicle Pet Barriers			8474
		Pet Strollers			6276
		Pet Steps & Ramps			6973
		Bird Supplies			:
			Bird Cage Accessories		7385
				Bird Cage Bird Baths	499954
				Bird Cage Food & Water Dishe	is 7386
			Bird Cages & Stands		4989
			Bird Ladders & Perches		4991
			Bird Toys		4992
			Bird Treats		4990
			Bird Food		4990
			Bird Gyms & Playstands		7398
		Pet Training Aids			505314
			Pet Training Pads		6846
			Pet Training Pad Holders		50530-
			Pet Training Clickers & Treat Dispensers		505313
			Pet Training Sprays & Solutions		50531

## **Google Product Taxonomy schema**



• PoolParty allows to import XLS/CSV, but one has to adhere to a certain target schema.

A	В	С	D	E	F	G
1	Animals & Pet Supplies					
3237	Animals & Pet Supplies	Live Animale	_			
2	Animals & Pet Supplies	Pet Supplies				
3	Animals & Pet Supplies	Pet Supplies	Bird Supplies			
7385	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories		
499954	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories	Bird Cage Bird Bat	hs
7386	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cage Accessories	Bird Cage Food &	
4989	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Cages & Stands	bit dage i dea a	
4990	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Food		
7398	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Gyms & Playstands		
4991	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Ladders & Perches		
4992	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Toys		
4993	Animals & Pet Supplies	Pet Supplies	Bird Supplies	Bird Treats		
4	Animals & Pet Supplies	Pet Supplies	Cat Supplies	bid nous		
5082	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Apparel		
4433	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Beds		
3367	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food		
543684	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food	Non-prescription C	at Food
543683	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Food	Prescription Cat Fo	
4997	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Fumiture	r leachption cat i c	iou -
500059	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Furniture Accessories		
4999	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter		
8069	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Box Liners		
7142	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Box Mats		
5000	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Litter Boxes		
5001	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Toys		
5002	Animals & Pet Supplies	Pet Supplies	Cat Supplies	Cat Treats		
5	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Gat fields		
5004	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Apparel		
4434	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Beds		
7372	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Diaper Pads & Liners		
499900	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Diapers		
3530	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food		
543682	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food	Non-prescription D	a Food
543681	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Food	Prescription Dog F	
5094	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Houses	Flescription bog F	000
7428	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Kennel & Run Accessories		
7274	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Kennels & Runs		
5010	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Toys		
8123	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Treadmills		
5011	Animals & Pet Supplies	Pet Supplies	Dog Supplies	Dog Treats		
5011	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Dog rieats		
505303	Animals & Pet Supplies	Pet Supplies Pet Supplies	Fish Supplies	Aquarium & Pond Tubing		
505303					1	
505307	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Air Stones & Diffuser	5	
500038	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Cleaning Supplies		
5 5019 7 5020	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Decor		
5020	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Filters		
	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Fish Nets		
9 5021	Animals & Pet Supplies	Pet Supplies	Fish Supplies	Aquarium Gravel & Substrates		

cheme	e concep	ot concept	concept		notatio
+	+	+			+
Animals & Pet Su	Pet Supplies				2
-	T Bt Ouppings	Pet Sunscreen			7396
		Vehicle Pet Barriers			8474
		Pet Strollers			6276
		Pet Steps & Ramps			6973
		Bird Supplies			3
			Bird Cage Accessories		7385
				Bird Cage Bird Baths	499954
				Bird Cage Food & Water Dishes	7386
			Bird Cages & Stands		4989
			Bird Ladders & Perches		4991
	_		Bird Toys		4992
			Bird Treats		4993
			Bird Food		4990
			Bird Gyms & Playstands		7398
		Pet Training Aids			505314
			Pet Training Pads		6846
			Pet Training Pad Holders		505304
			Pet Training Clickers & Treat Dispensers		505313
			Pet Training Sprays & Solutions		50531

## **OntotextRefine for data wrangling**

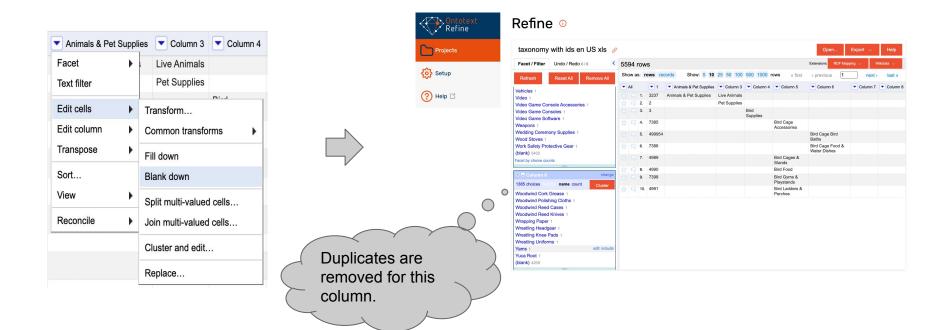


<ul> <li>Animals &amp; Per</li> </ul>	Supplies Column 3 Column
Facet	Text facet
Text filter	Numeric facet
Edit cells	Fimeline facet
Edit column	Scatterplot facet
Transpose	Custom text facet
Sort	Custom numeric facet
View	Customized facets
Reconcile	•

Refrest       Rearrow AV         Refrest       Rearrow AV         Refrest       Rearrow AV         Refrest       Rearrow AV         Refrest       Calcumate & Pet Supplies       Column 4       Column 5       Column 6       Column 7       Column 7         Refrest       Refrest       Calcum 7       Column 4       Column 5       Column 6       Column 7       Column 7         Refrest       Refrest <t< th=""><th></th><th>taxonomy with ids en</th><th>US xls 🔗</th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th>Open</th><th>Export ~</th><th>Help</th></t<>		taxonomy with ids en	US xls 🔗								Open	Export ~	Help
Help C   Help C Kertech Kertech Caluma Counce   Help C Kertech Caluma Counce Caluma Counce   Supples Caluma Counce Caluma Counce Caluma Counce   21 choices namais & Pet Supples Lev Animais & Pet Supples Lev Animais & Pet Supples   21 choices namais & Pet Supples Caluma Counce Caluma Counce   21 choices namais & Pet Supples Lev Animais & Pet Supples Lev Animais & Pet Supples   21 choices namais & Pet Supples Lev Animais & Pet Supples Pet Supples   21 choices namais & Pet Supples Lev Animais & Pet Supples Pet Supples   21 choices 2 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 3237 Animais & Pet Supples Pet Supples   21 choices 1 324 Animais & Pet Supples Pet Supples   21 choices 1 324 Animais & Pet Supples Pet Su	Setup	Facet / Filter Undo / Redo 0	/0	124	mate	hing ro	OWS (5594 total)				Extensions RDF Maj	oping 🗸 🛛 V	Nkidata 🗸
In the stand		Refresh Reset All	Remove All	Show	v as:	ows rec	cords Show: 5 10	25 50 100	500 1000	rows « first	<previous 1<="" th=""><th>next &gt;</th><th>last »</th></previous>	next >	last »
Supplies       1.3237       Animals & Pet Supplies       Lev Animals       Med Animals <th>Help 🖸</th> <th>2 = Animala 8 Pot</th> <th>anne invert reset</th> <th>💌 Al</th> <th></th> <th>• 1</th> <th><ul> <li>Animals &amp; Pet Supplies</li> </ul></th> <th>Column 3</th> <th>Column 4</th> <th>Column 5</th> <th>Column 6</th> <th>Column 7</th> <th>Column</th>	Help 🖸	2 = Animala 8 Pot	anne invert reset	💌 Al		• 1	<ul> <li>Animals &amp; Pet Supplies</li> </ul>	Column 3	Column 4	Column 5	Column 6	Column 7	Column
			ingo invort robot				Animals & Pet Supplies	Live Animals					
Animals & Pet Supplies Part Supplies 124 exclute Apparel & Accessories 240 Arts & Entertainment 1500 Baby & Toddier 97 Business & Industrial 224 Cameras & Optics 104 Electronics 418 Food, Bevergages & Tobacco 364 Furniture 121 Hardware 522 Health & Beauty 346 X C S M995 Animals & Pet Supplies Animals & Pet Supplies Pet Supplies Pird Cage Bird Cage Bird Bird Cage Bird Bird Cage Bird Cage Bird Cage Bird B		21 choices Sort by: name count	Cluster										
Arts & Entertainment 500       Barly & Toddler 87       Supples       Supples       Normals & Pet Supples       Pet Supples       Bird Cape       Bird C		Animals & Pet Supplies 124	the second se		3.	3	Animals & Pet Supplies	Pet Supplies					
Image: Solution S					7 4.	7385	Animals & Pet Supplies	Pet Supplies					
Business & Industrial 2243 Cameras & Optios 104 Electronics 418 Food, Beverges & Tobacco 364 Furniture 121 Hardware 522 Health & Beauty 366	Þ	Baby & Toddler 87			5.	499954	Animals & Pet Supplies	Pet Supplies					
Food, Beverages & Tobacco 364     Fundure 121       Humiture 121     Animats & Pet Supplies       Health & Beauty 346       Image: Supplies       Image: Supplies </td <td></td> <td></td> <td><b>7</b> 6.</td> <td>7386</td> <td>Animals &amp; Pet Supplies</td> <td>Pet Supplies</td> <td>Bird</td> <td>Bird Cage</td> <td>Bird Cage Food &amp;</td> <td></td> <td></td>					<b>7</b> 6.	7386	Animals & Pet Supplies	Pet Supplies	Bird	Bird Cage	Bird Cage Food &		
Hardware 522     Supples     Supples       Health & Beauty 346     Part Supples     Bird     Bird Syns & Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Bird Gyns & Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Bird Gyns & Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Bird Gyns & Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Pet Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Pet Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Pet Supples       Image: Supples     Pet Supples     Bird Syns & Supples     Pet Supples					7.	4989	Animals & Pet Supplies	Pet Supplies					
Restitutivation document     Restitutivation     Restitutivation <td></td> <td></td> <td></td> <td>7 8.</td> <td>4990</td> <td>Animals &amp; Pet Supplies</td> <td>Pet Supplies</td> <td></td> <td>Bird Food</td> <td></td> <td></td> <td></td>					7 8.	4990	Animals & Pet Supplies	Pet Supplies		Bird Food			
X = Columni 3     Change       2 choices Sort is name count     Cluster       Live Animals 1     Cluster		Health & Beauty 346			9.	7398	Animals & Pet Supplies	Pet Supplies	Bird				
Live Animals 1					7 10.	4991	Animals & Pet Supplies	Pet Supplies					
		2 choices Sort by: name count	Cluster										
		Live Animals 1											
Pet Supplies 123		Pet Supplies 123											

## **OntotextRefine for duplicate removal**





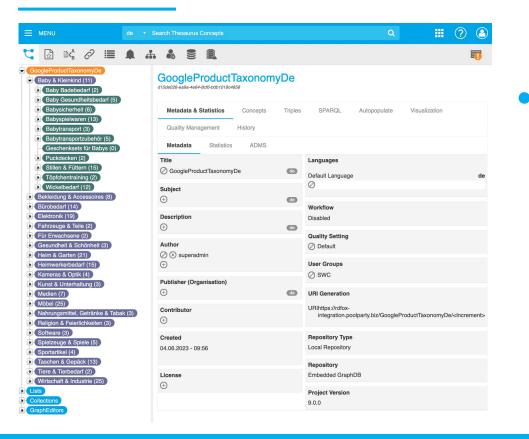
## After import of "en" version



≡ MENU en •	Search Thesaurus Concepts	۹ 🗰 🧿 🔕				
5 🕼 🕼 🧷 🗮 🌲 🞜		<b>1</b>				
CooglePidducTixxonomyEn     Animals & Pot Supplies (2)     Live Animals (0)     Pot Supplies (46)     Apparel & Accessories (6)     Arts & Entertainment (3)	Add to Blacklist     a Add to ExactMatch     Live Animals     fitue.itsdise.integration.coclearly.biolScoplePtodecTlassocaryEis/1	+ Create Narrower Concept				
Baby & Toddler (11) Business & Industrial (25) Cameras & Optics (4)	Details Notes Documents Linked Data	Triples Visualization Quality Management				
Electronics (19)     Food, Beverages & Tobacco (3)	SKOS <u>•</u> +					
Furniture (25)     Hardware (15)     Health & Beauty (3)	Broader Concepts	Preferred Label Ø Live Animals				
Home & Garden (21) Luggage & Bags (13)	Narrower Concepts	Alternative Labels				
Mature (2) Media (7) Office Supplies (14)	Related Concepts	Hidden Labels ⊕				
Religious & Ceremonial (3)     Software (3)     Sporting Goods (4)     Toys & Games (5)	Top Concept of Concept Schemes Animals & Pet Supplies Ø	Notation ⊘ ⊗ 3237 ⊙				
Vehicles & Parts (2)	Exact Matching Concepts	Scope Notes ⊕				
GraphEditors	Close Matching Concepts	Example (+)				
	Broader Matching Concepts	Definitions (-)				
	Narrower Matching Concepts	$\bigcirc$				
	Related Matching Concepts					
		on is mapped s:notation				

- Graph data model in RDF is flexible to do data merging (unlike tabular data).
- Use skos:notation as an identifier to merge "de" version.

## After import of "de" version



- Two projects "en" & "de" live in different repositories, we have to:
  - use query federation, or
  - use a remote GraphDB



## Creating repository in external GraphDB





Create

Cancel

© Semantic Web Company

# Exporting..

• Exporting concepts for both "en" and "de" versions results with two named graphs that holds all the respective taxonomy data.

		×
Export Options		
Template	Select Template	Ŧ
Action	Export to GoogleProductTaxonomy	•
Graph URI	https://rdfox-integration.poolparty.biz/GoogleProductTaxonc	omyDe/thesauru
Format	Trig	•
Thesaurus Data	Concepts X	
Exclude Concept Data	Select Items	•
Include Custom Data	Select Items	×
	Show More	
	Export	Cancel



Note in GraphDB: For a given query x without specifying a named graph results in default graph [7]  $q(x) = q(x, G1 \cup G2 \dots)$ In our case:  $q(x) = q(x, G_{en} \cup G_{de})$ 





• Note that we use skos:notation as *identifiers* to do the merging.

```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
```

## After RDF merging



	? 3
C 🖸 🖏 🤣 📰 🌲 🌲 🥃 🌉	<b>I</b>
CooglePhodueTaxonomy     CooglePhodueTaxo	elete Concept
Optimized Statistication         Details         Documents         Linked Data         Triples         Visualization           Comments (10)         Outlity Management         History         Outlity Management         History	
Food, Beverages & Tobacco (3)     SKOS ++	
Iterativane (15)         Broader Concepts         Preferred Label           • (Health & Beauly (5)         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •         •	en
Luggage & Bags (13)     Narrower Concepts     Matrixe (2)     Concepts     Metrix (2)     Concepts     Metrix (2)     Concepts	0
Office Supplies (14)     Related Concepts     Office Supplies (14)     Related Concepts     Hidden Labels     Office (15)	
Sponting Goods (4)         Top Concept of Concept Schemes           Toys & Games (5)         Animals & Pet Supplies         Notation           Vehicles & Parts (2)         Ø         Ø         3237	
Collections     Exact Matching Concepts       CalphEditors     Ø @       Scope Notes       O	
Close Matching Concepts	
Broader Matching Concepts	
Narrower Matching Concepts Ø ☺	
Related Matching Concepts Ø ☺	

Add to Eliacklist • Add to ExactMatch		+ Create Narrower Concept.	elete Concep
tere Annu Construction accelerative Also Stopphe Product Taxonomy Env1 Details Notes Documents Linked Data Triple	ss Visualization Quality Management History		
As Object As Obj			
Predicate 1	Object	Graph	
http://purl.org/dc/terms/created	2023-06-04T07:49:18.566Z (##035667Tme)	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	Î
http://purl.org/do/terms/oreator	https://rdfox-integration.pooloarty.biz/user/superadmin	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	î
http://www.w3.org/1999/02/22-rdf-syntax-nsi/type	http://www.w3.org/2004/02/skos/core#Concept	https://tdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	Ŧ
http://www.w3.org/2004/02/skos/core#notation	3237 (pointer)	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	
http://www.w3.org/2004/02/skos/core#prefLabel	Lebende Tiere do	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	т
http://www.w3.org/2004/02/skos/core#prefLabel	Live Animals (m)	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	î.
http://www.w3.org/2004/02/skos/core#topConceptOf	https://rdlox-integration.poolparty.biz/GoogleProductTaxonomyEn/D	https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus	î

## After RDF merging (cnt.)



https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus

https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus

https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus

https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesauru

≡ MENU en • S	earch Thesaurus Concepts	۹ 🗰 🥐 🙆				
🤁 🗟 🖓 🗐 🌲 🕯	h & S R	<b>I</b>				
GoogleProducTaxonomy Animals & Pet Supplies (2) Live Animals (0) P Pet Supplies (46) Apparel & Accessories (8) Aptaret & Accessories (8)	Add to Blacklist     Add to ExactMitteh     Live Animals     Masufistic Integration assessments And Geogle Decker Timenemy Engl	+ Create Narrower Concept				
Baby & Toddler (11)     Business & Industrial (25)     Cameras & Optics (4)     Electronics (19)     Food, Beverages & Tobacco (3)	Details Notes Documents Linked Quality Management History	d Data Triples Visualization				
Furniture (25)	SKOS • +					
Hardware (15) Health & Beauty (3) Home & Garden (21)	Broader Concepts	Preferred Label Ø Live Animals				
Luggage & Bags (13) Mature (2)	Narrower Concepts	Alternative Labels				
Media (7)     Office Supplies (14)     Religious & Ceremonial (3)     Software (3)	Related Concepts	↔ Hidden Labels ↔				
Sporting Goods (4)	Top Concept of Concept Schemes	$\oplus$				
Toys & Games (5) Vehicles & Parts (2)	Animals & Pet Supplies	Notation				
Lists Collections GraphEditors	Exact Matching Concepts	Generation Scope Notes				
	Close Matching Concepts	⊕ Example				
	Broader Matching Concepts	⊕				
	Ø 🛛	Definitions (+)				
	Narrower Matching Concepts	0				
	Related Matching Concepts					

#### https://.../GoogleProductTaxonomyEn/1

3237 assisting

Lebende Tiere

Live Animals (0)

http://www.w3.org/2004/02/skos/core#notation

http://www.w3.org/2004/02/skos/core#prefLabel

http://www.w3.org/2004/02/skos/core#prefLabel

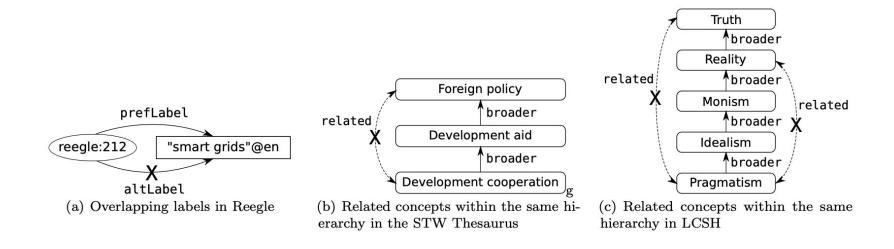
http://www.w3.org/2004/02/skos/core#topConceptOf

VS **3237** (skos:notation) • Mapping done for us, but we didn't have much control. + Add to Blacklist + Add to ExactMatch Live Animals Create Narrower Concept... 👕 Delete Concept Linked Data Triples Visualization Quality Management Histon As Subject biect: htt Predicate 1 Object Graph http://purl.org/dc/terms/created 2023-06-04T07:49:18.566Z https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesauru http://purl.org/dc/terms/creator https://rdfox-integration.por https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus http://www.w3.org/1999/02/22-rdf-syntax-ns#type http://www.w3.org/2004/02/skos/core#Cor https://rdfox-integration.poolparty.biz/GoogleProductTaxonomy/thesaurus

https://rdfox-integration.poolparty.biz/GoogleProductTaxonomyEn

## **Ensuring SKOS data quality**





Osma Suominen, Christian Mader: Assessing and Improving the Quality of SKOS Vocabularies, 2014

## Data validator for SKOS compliance



:x a :Concept , :A, :B . :A owl:disjointWith :B . :x a :Concept, :ConceptScheme. GoogleProductTaxonomy . :x :p1 ""; :p2 "". :p1 rdfs:domain :A . :p2 rdfs:domain :B . da3/7538-06ef-4214-acd3-8095bb1d148d :A owl:disjointWith :B. Metadata & Statistics SPARQL Autopopulate Quality Managemen Quality Report Data Validator :x dcterms:created "2022-11-23T21:24:38.716Z"^^xsd:dateTime Cameras & Optics (4) Electronics (19) :x a ConceptScheme . Food, Beverages & Tobacco (3) During data validation, you will not be able to edit the thesaurus, Project Linking and Ontology I
 You will be able edit data in these places again after you have saved imported or disparded the Furniture (25) Hardware (15) .....: :y skos:broader :z. :z :skos:topConceptOf :x . Health & Beauty (3) Results (Last Validation on Thu Jun 22 07:25:25 CEST 2023) y skos:broader :z. :z skos:narrower :y : Home & Garden (21) Luppage & Bags (13) No disjoint classes found. Description Mature (2) Media (7) No conflicting properties found. Description Office Supplies (14) Religious & Ceremonial (3) All necessary date and time resources are correct. Descript ..... :x dc:title/rdfs:label "x"@en. Software (3) Concept scheme found. Description Sporting Goods (4) Toys & Games (5) Hierarchical consistency verified. Description Vehicles & Parts (2) .....:y skos:prefLabel "y" . No missing inverse relations found. Description .....y skos:prefLabel "y1"@en , "y2"@en . No deprecated resources found. Description All concept schemes have a title. Description ..... :y skos:prefLabel ""@en . All labels have language tags. Description All concepts have a preferred label (default language) Description :y skos:prefLabel "y"@en . :z skos:prefLabel "y"@en. No empty labels found. Description ..... :y skos:prefLabel "y"@en, "y"@de . All preferred labels are unique per language. Description All concept labels are disjoint. Description .....: :y skos:notation :id123 . skos:notation rdfs:range :Literal . All attribute values are literals. Description SKOS collections have the correct "rdf:first" range. Description Domain and range for SKOS properties satisfied. Description No multiple usage of unique properties. Description .....::y:id:id123.:id rdfs:range:Literal; rdfs:domain:Person. Domain and range for custom scheme properties satisfied. Description :v :id "123", "1234" . :id a :FunctionalProperty .



- PoolParty allowed us to import XLS and created RDF;
- The mapping from XLS -> RDF is done via *implicit mappings* by PoolParty.
- As we didn't have much control of it, we ended up with inconsistent URIs vs skos: notation.
- Despite we can repair them, **URIs** have to be considered **stable** and only change them **before ingestion**.
- We can alleviate this with *explicit* and *declarative mappings*.

## **Extracting concepts from text (NLP)**



Display Debug Information

## "The **guitar** is a fretted **musical instrument** that typically has six strings."



Concepts					Display Debug Information
Concept Preferred Label	Normalized Score	Corpora Score	Transitive Broaders	Transitive Broader Top Concepts	Related Concepts
Guitars	100.0				
Musical Instruments	54.0				
Shadow Concepts					
Concept Preferred Label	Normalized Score	Corpora Score	Transitive Broaders	Transitive Broader Top Concepts	Related Concepts
Free Terms					
Term		Normalized Score	Corpora	Score	Frequency in Document
musical instrument that typically		100.0			

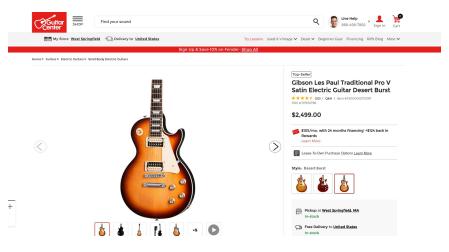
	Hormanizou oboro	eeipona eeene	riequency in Decument
musical instrument that typically	100.0		1
fretted musical	59.0		1
fretted	59.0		1
typically has six strings	53.0		1
musical	50.0		1
instrument that typically	41.0		1
instrument	41.0		1
typically	27.0		1
strings	11.0		1

Powered by PoolParty Extractor

## **Use case: GPT Semantic Search**



- Let us quickly create a semantic search on guitars based on GPT..
- We take *unstructured text* of the guitar's description and *annotate it* using PP Extractor;
- We directly store the results to the remote GraphDB using annotate/store API.



## Guitars with annotated data in RDF



S GoogleProductiaxonomy ~ (2) anmetia

ce3975ff-e70c-4a59-89e2-

http://schema.semantic-web.at/ppx/frequency

http://schema.semantic-web.at/ppx/referenceThesaurus

da3f7538-06ef-4214-acd3-8095bb1d148d

http://schema.semantic-web.at/ppx/score

Type

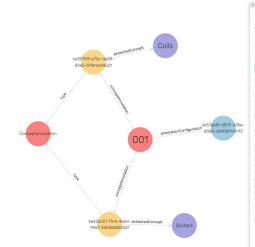
cb57a65f-d879-435e-

8-c840b91d6

RDF rank:

x xsd:integer

71.0 xsd:double





Main node containing description of the guitar.

Tagging event with score and frequency

The tagging event can be written in a more compact

001

form using RDF\*.

2967-34h15h80

:03975ff-070r

## Guitars with annotated data in RDF\*

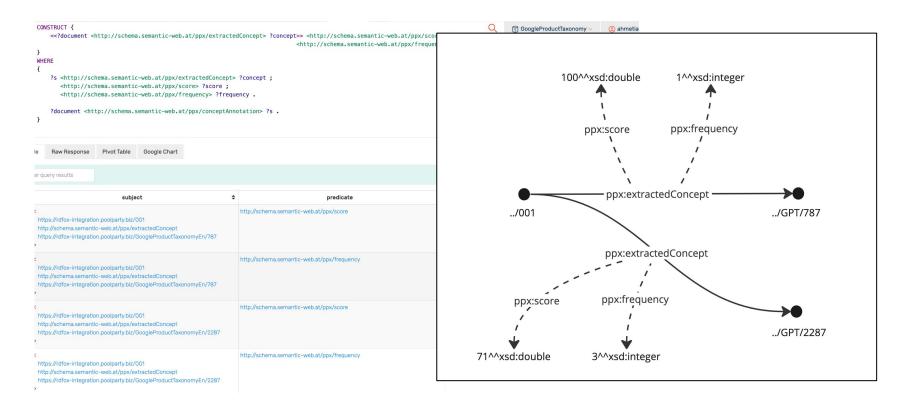


	subject		predicate	\$		object	
er query results					Showing	results from 1 to 22 of 22. Query took	0.2s, moment
e Raw Response	Pivot Table Google Chart					Download	d as 🗸 🛝 🛝
}							R keyboard sh
?document <ht< td=""><td>ttp://schema.semantic-web.at</td><td>t/ppx/conceptAnnotation&gt; ?s .</td><td></td><td></td><td></td><th></th><td>(</td></ht<>	ttp://schema.semantic-web.at	t/ppx/conceptAnnotation> ?s .					(
<http: sc<="" td=""><td>chema.semantic-web.at/ppx/ex chema.semantic-web.at/ppx/so chema.semantic-web.at/ppx/fo</td><td>core&gt; ?score ;</td><td></td><td></td><td></td><th></th><td></td></http:>	chema.semantic-web.at/ppx/ex chema.semantic-web.at/ppx/so chema.semantic-web.at/ppx/fo	core> ?score ;					
{							
} WHERE							
	<pre>circ(p.//schema.semaircic-web.</pre>	at/ppx/extractedconcept> reone	<pre><http: frequ<="" ppx="" pre="" schema.semantic-web.at=""></http:></pre>		y.		
CONSTRUCT {	chttn://schema_semantic_web	at/onv/extractedConcent> ?conc	<pre>cept&gt;&gt; <http: ppx="" pre="" sc<="" schema.semantic-web.at=""></http:></pre>	ores ?score	Q	GoogleProductTaxonomy ~	(2) ahmet

subject	predicate	object
<pre>c https://rdfox-integration.poolparty.biz/001 http://schema.semantic-web.at/ppx/extractedConcept https://rdfox-integration.poolparty.biz/GoogleProductTaxonomyEn/787 </pre>	http://schema.semantic-web.at/ppx/score	*100.0** xsd.double
c https://rdfox-integration.poolparty.biz/001 http://schema.semantic-web.at/ppx/extractedConcept https://rdfox-integration.poolparty.biz/GoogleProductTaxonomyEn/787	http://schema.semantic-web.at/ppx/frequency	*****ksdinteger
c https://rdfox-integration.poolparty.biz/001 http://schema.semantic-web.at/ppx/extractedConcept https://rdfox-integration.poolparty.biz/GoogleProductTaxonomyEn/2287	http://schema.semantic-web.at/ppx/score	*71.0** <sup>°</sup> xii double
<pre>c https://rdfox-integration.poolparty.biz/001 http://schema.semantic-web.at/ppx/extractedConcept https://rdfox-integration.poolparty.biz/GoogleProductTaxonomyEn/2287 </pre>	http://schema.semantic-web.at/ppx/frequency	*3*" "kunneger

## Guitars with annotated data in RDF\* (viz.)





## **RDF\*** vs other modelling approaches



Modeling approach	Total statements	Loading time (min)	Repository image size (MB)
Standard reification	391,652,270	52.4	36,768
N-ary relations	334,571,877	50.6	34,519
Named graphs	277,478,521	56	35,146
RDF-star	220,375,702	34	22,465

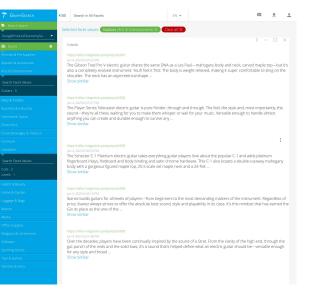
https://graphdb.ontotext.com/documentation/10.0/devhub/rdf-sparql-star.html

Daniel Hernández et al. Reifying RDF: What works well with wikidata?, 2015.

## **GPT Semantic Search**

- Semantic Search as a "smoke test" to check how the taxonomy fits your data;
- Not many hits with GPT, but Guitars are domain-specific;
- GPT can be still used as an **upper taxonomy**;

We can use the corpus of data to enrich the taxonomy as a *feedback loop for extraction and search*.





## **Corpus analysis - Composite AI with statistics**



							II 🥐 🤅
) № 2 🗏 🌲 📥							
us als & Pet Supplies (2) rel & Accessories (8)	GuitarCenter corpus:2056/06/2-65/1-4/3d-92864-64/2058-06/07						Corpus Searc
& Entertainment (3) rent Tickets (0)	Metadata & Statistics Extracted Concepts	Extracted Terms Corpus Documents					
obbies & Creative Arts (8)	Search Terms		WSI Filter				
Arts & Crafts (6) Collectibles (11)	Search		All		-		
Homebrewing & Winemaking Supplies (4) Juggling (0)	Search Reset						
Magic & Novellies (0) Model Making (4) Musical Instrument & Orchestra Accessories (14)	Extracted Terms				Select A	Add Candidate Concepts A	dd to Blacklist Export Documents
Musical Instruments (8) Accordions & Concertinas (0)	Term	Relevance	CTS	- MIS	Frequency		
Bagpipes (0) Brass instruments (6)	Accessories	1495.44	186.5	0	41131		• 2 0
Electronic Musical Instruments (4) Percussion (11)	Compare	582.14	146	0	25764		• 0 0
Pianos (0) String Instruments (6)	Stands	455.85	62.167	0	35300		• 0 0
Cellos (0) Guitars (0) Harps (0)	Price	223.59	46.625	0	33481		• 2 🛇
Upright Basses (0) Violes (0)	Drums	193.54	31.083	0	25640		• @ 🛇
Violins (0) Waadwinds (14)	inCheckoutPromo	257.32	27.241	0	12657		• 2 🛇
arty & Celebration (4) & Toddler (11)	New	102.98	26.643	0	25057		• @ 🛇
iess & Industrial (25) eras & Optics (4)	Bags	57.01	23.313	0	24963		• 2 🛇
ronics (19) , Beverages & Tobacco (3) ture (25)	Bass	132.37	20.722	0	24554		• @ 🛇
ware (15) h & Beauty (3)	Sybersound	13.7	20.433	0	46		• @ 🛇
e & Garden (21) age & Bags (13)						1 2 3 4	5 6 7 8 9 10 Next Last
re (2) B (7)							

© Semantic Web Company

Religious & Cerer

## **Corpus analysis - Composite AI with statistics**



T PROJECT CORPORA TOOLS ADVANCED						۵ 🛛 🖩 🖉 ک	7 Upload Documents
5 🖸 🖬 🖉 🔳 🖉 🖉	R.						Upload Documents Paste Text Crawl Website Crawl RSS Feed Crawl DBpedia
Thiskaurus     Animala & Pet Supplies (2)     Apparel & Accessories (0)	GuitarCenter					R Corpus Search	Provide UFL: https://www.gultarcenter.com/Fender/Player-Stratocaster-HSS-Pau-Feno-Fingerboard-Umited-Edition-Electric-Gultar-Ca
Artis & Enterhairement (3)     Event Trokets (0)	Metadata & Statistics Extracted Conce	epts Extracted Terms Corpus Documents					Depth: 1
Hobbies & Creative Arts (8)     Arts & Creatis (8)	Search Terms		WSI Filter				Enforce Corpus Language:
Collectibles (11)     Hometrewing & Whemaking Supplies (4)     Agging (0)	Search Reset						Emotos Corpos Language. 🛛 🔽
(Magio & Novelies (0)     (Magio & Novelies (0)     (Magio & Novelies (0)     (Magio & Instrument & Orchestra Accessories (14)	Extracted Terms				Select All	Deselvot All Add Candidate Concepts Add to Blacklist Export Documents	Website successfully added to corpus.
Musical Instrument 8 Chorestin Accessiones [14]	Term	Relevance	CTS	- MIS	Frequency		Crewt
	Accessories	1495.44	186.5	0	41131	• @ 🛇	Ctose
Percussion (11)	Company	582.14	146	0	25764	• 2 0	0000
	Stands	455.85	62.167	0	35300	• @ 🛇	
- Guitare (0) - Guitare (0)	Price	223.59	46.625	0	33481	• d 🛇	
Upright Basses (0) Vicias (3)	Drams	193.54	31.083	ō	25640	• 2 🛇	
Viciins (D) Woodwinds (14)	in CheckoutPromo	257.32	27.241	0	12657	• 2 🛇	
Party & Celebration (4)     Baby & Todder (11)     Business & Industrial (25)	Network	102.98	26.643	0	25057	• @ 🛇	
Carneras & Optics (4)	Bags	57.01	23.313	0	24963	• @ 🛇	
Food, Beverages & Tobacco (3)     Fumilure (25)     Handware (15)	Bass	132.37	20.722	0	24554	• 2 🛇	
Nardware (15)     Health & Beauty (3)     Homo & Gasten (21)	Sybersound	13.7	20.433	0	46	• @ 🛇	
Luggage & Bags (13)						1 2 3 4 5 6 7 8 9 10 Nast Last	
Mature (2)							

- Relevance: The Mutual Information Score, the Content Term Score and the term frequency were combined into one score that gives an overall relevance that is normally a good starting point for going through the list of extracted terms.
- Mutual Information Score (MIS): Mutual information (MI) provides information about dependency of variables and can be used to estimate if two or more consecutive words in a text should be considered a compound term that is formed by those words. The idea is that if words are independent that they will occur together just by chance. On the other hand, if they are observed together more often than expected, they are dependent and are candidates for terms. This score ranks multi term phrases higher.
- Content Term Score (CTS): Content terms are enriched in documents where they appear, which means they are not the most frequent terms in the document set. But when they occur in a document they tend to occur very often, which indicates their importance for defining the content of the document. This score ranks single term phrases higher.
- Frequency: The total number of occurrences of a term in the corpus.

## Adding recommendations



			×		
$\leftrightarrow$ $\rightarrow$					
No image available	DESCRIPTION The Gibson Trad Pro V electric guitar shares the same DNA as a Les Paul—mahogany body and neck, canved maple top- instrument. You'll feel if from: The body is weight releved, maling it super comfortable to sling on the shoulder. The neck like a glow. Thenks to the compound radius reservation of freboart, you'll fiel is easier to solon higher registers. And that too, Gibson loads the Po V with dual Tradhucker pickups. The neck position humbucker is underwound for a varm, org overwound, increasing the output. Both use Anico Y wagnets. Tweeky you trons with advanced electronics A put-typuil your choice via a dip switch, perfect if your tone pool is filled with deep British blues. The Trad Pro V a tremendou comes with a hardshell case. Arts & Entretranment Haroware	k has an asymmetrical shape that fits 's where you'll hear the difference, ganic sound; the bridge is performs a coil split or coil tap—it's and outer coils. The rhythm tone	See Also Thesaurus Samlaru The aurus Samlaru The		
Recommended General Record		NSERT DAT GRAPH < :001 <h< th=""><th>.&gt; {</th><th>antic-web.at/recommends&gt;</th><th>&gt; :005</th></h<>	.> {	antic-web.at/recommends>	> :005

### Adding recommendations using a SPARQL rule



			×	
$\leftrightarrow$ $\rightarrow$				
No image available	Dischiption The Gibson Trad Pro V electric guitar shares the same DNA as a Les Paul—mahogany body and neck, carved mi instrument. You'll feel it first: The body is weight relieved, making it super comfortable to sling on the shoulder. I ke a glove. Thanks to the compound radius rosewood fretboard, you'll find it easier to solo in higher registers , too. Gibson loads the Pro V with dual Tradbucker pricipats. The neck position humbucker is underwound for a w overwound, increasing the output. Both use Anico V magnets. Tweak your tones with advanced electronics. A p your choice via a dip switch in the control cavity, to coll-split mode, the treble tone push/pull toggles between th push/pull activates a phase switch, perfect if your tone pool is filled with deep British blues. The Trad Pro V a tre comes with a hardshel case. Arts & ENTERTAINMENT HADWARE	The neck has an asymmetrical shape that fits And that's where you'll hear the difference, arm, organic sound; the bridge is ush/pull performs a coil split or coil tap—it's e inner and outer coils. The rhythm tone	See Also Thesaurus Similarity The Content The Content	
	Rule head	INSERT { GRAPH < ?s <http: commonst<="" commonstance.com="" initial="" th="" www.commonstance.commonstance.com=""><th>-</th><th>tic-web.at/recommends&gt; ?o</th></http:>	-	tic-web.at/recommends> ?o
Recommended General Reco	Rule body	} WHERE { ## in: ## bi: ## bi:	nd ?s	c here ##
		}		

#### Part III: The secrets of cooking









- Each of the rule is of the form Head :- Body
- Each variable in the head of each rule must appear in the body of the rule
- "Negation as failure" using not syntax

#### **Datalog example**



ABOX: :ann a :Person ; :worksFor :upc .

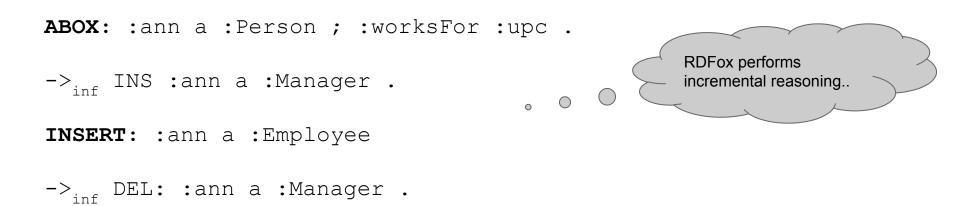
```
->_{inf} INS :ann a :Manager .
```

**INSERT:** : ann a : Employee

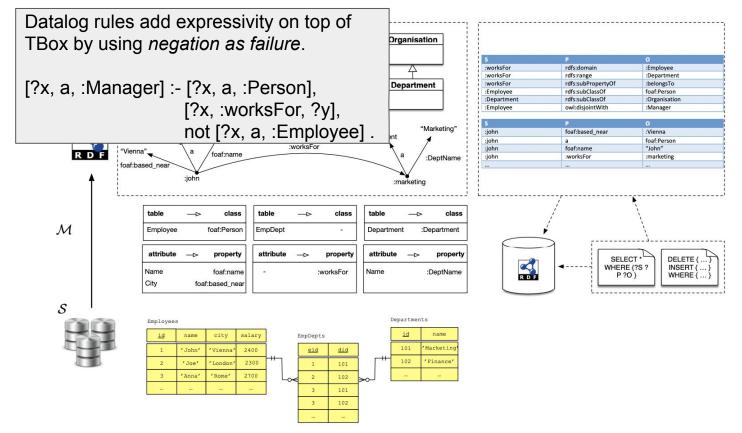
```
->_{inf} DEL: :ann a :Manager .
```

#### Datalog example (cnt.)





# Adding Datalog rules for "common sense" reasoning poolparty.





## Use case: Data Validation with Datalog & SHACL

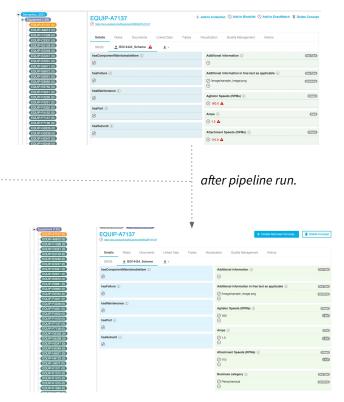
Add to Collection 🛇 Add to Blacklist 🛇 Add to ExactMatch 👕 Delete Concep FOUIP-A713 QUIP-A9812 ( Ensure consistent project data regardless of the source If inconsistent, repair the data. 115 A ⊙ 110.0 A Minimal change <urn:shapes-inferred>(?s, rdf:type, sh:PropertyShape), Datalog rules in RDFox principle. <urn:shapes-inferred>(?s, sh:targetClass, skos:Concept), DELETE { GRAPH < https://el-capitan.poolparty.biz/Semantics 2021 esaurus> <urn:shapes-inferred>(?s, sh:path, ?p), ?focusNode ?property ?value }} Repair using <urn:shapes-inferred>(?s, sh:datatype, ?d) :-SPARQL/Update INSERT { GRAPH <https://el-capitan.poolparty.biz/Semantics 2021/thesaurus> { ?focusNode ?property ?new }} poc:ISO14224Ontology(?p, rdf:type, owl:DatatypeProperty), WHERE { poc:ISO14224Ontology(?p, rdfs:range, ?d), GRAPH <urn:report> { rdfox:SKOLEM("Shape", ?d, ?p, ?s). ?s sh:sourceConstraintComponent sh:DatatypeConstraintComponent . ?s sh:focusNode?focusNode. Create shapes from Ontology axioms ?s sh:resultPath ?property. SHACL validation ?s sh:value ?value . } swcs:DatatypeShape a sh:PropertyShape; report GRAPH <https://poc.poolparty.biz/ISO14224Ontology> { ?property rdfs:range sh:targetClass skos:Concept; ?range. } sh:path <https://poc.poolparty.biz/ISO14224Ontology/fuel>; BIND(STRDT(?value,?range) as ?new). sh:datatype <http://www.w3.org/2001/XMLSchema#boolean>.

#### **Data Validation & Repairs**



• If inconsistent, repair the data.

Validate and repair pipeli	ne' ↓			
Pesigner				
ailable DPUs 🔨	-			
Private DPU Templates				
e to Filter Tree				
ictors				
les Download		SPARQL Endpoint Loader		SPARQL Endpoint Loader
		Validate	run_after	Repair
driveExtractor			Iun_atter	Repair
TTP API Request				
elational Database		.*		
Ik RDF Link Discovery		.:		
PARQL Endpoint				
lers				
ers				
ers) 7: SMRCL Dispoint Loader Debil				۰×
CETS				* ×
OTS				* ×
OFS T SWARL Exception Loader Detail Name: StrateC. Encoder Loader Parent: StrateC. Encoder Loader				* *
OTS T SMAC, Corport Landr Unit News: SPACE, Deport Land Partit: SPACE, Deport Land Description Volcar		Una Templek Darliganten		
T SMALL Evenes Leave Insu     News: SMALL Evenes Leave Insu     News: SMALL Evenes Leave Insu     News: SMALL Evenes Leave     Nuclear     SMALL Evenes Leave Insu     SMALL     SMALL EVENUE Leave Insu				
OTS T SMAC, Corport Landr Unit News: SPACE, Deport Land Partit: SPACE, Deport Land Description Volcar		Chen Terripi en Configuration EMPEGA Engigent Inflatestimultachen una se sessional	Fas 1	
OFFS     OFFSC Droken Land Deal     Nove: String Droken Land     Pert: String Droken Land     Pert: String Droken Land     Pert Configuration     Voides     Pert Configuration     Nove:     Pert Configuration     Nove:		SPAROL Endpoint * idatations/customer-use-case/sparol Username *	12110 Password *	••
The second	e Dia Catan Description	SFNACL Endpoint " Edutations/custome Case-Spacej Uservane " admin	12110	* X
CONTRACT CONTRACT OF CONTRACT	ir V Una Caston Discription	SP4AD4, Endpoint * distantenet customer case taxes lipangl Usernese * admin genetity Target Corpti	12110 Password *	**
The second	or Use Custor Decorptor State Activitiation Tot File Function	SFNACL Endpoint " Edutations/custome Case-Spacej Uservane " admin	12110 Password *	* *
Compare Law Part Compare Law Part Part Part Part Part Part Part Part	er Cater Description State Aufternation State Fair Faunat There incomests	SPMDAL Endpoint * Statustices/castoner.cas-topang Useraes * afrite Gopoly Target Croph Graph LEV	12110 Password *	* *
Alter      Alter     Alter      Alter      Alter      Alter      Alter      Alter	er Cale Custor Deception Cale Custor Decept	SPMDAL Endpoint * Statustices/castoner.cas-topang Useraes * afrite Gopoly Target Croph Graph LEV	12110 Password *	* *
Provide Language can be for the second	er Cale Custor Deception Cale Custor Decept	STATUS_Colgoine * Samananatustamerese establigation Marrene * and Bearly Turget Courts Courts Derskutz	12110 Password *	**
Provide Language can be for the second	er California Description California Descrip	STATUS_Colgoine * Samananatustamerese establigation Marrene * and Bearly Turget Courts Courts Derskutz	12110 Password *	* *
Provide Language can be for the second	er California Description California Descrip	STATUS_Colgoine * Samananatustamerana asahagasi Marrene * ana @ Bandh Turget Count Grand Turget Count	12110 Password *	* * *





#### **RDFox recommender**



poolparty		
	How can i combine KGs and semantic reasoning as part of my Al strategy into my machine learning to obtain an explainable Al and improve my data quality	
	Example 1 Example 2 Example 3	
	GET RECOMMENDATION	

nowledge graphs <b>x</b> Reasoning <b>x</b> AI strategy <b>x</b> Machine learning <b>x</b> Explainable AI <b>x</b> Data quality <b>x</b>	$\times   \vee$
lecommended Content	^
How are knowledge graphs related to artificial intelligence?	^
Score: 0.39566260121140084 Knowledge graphs are at the core of semantic artificial intelligence. Semantic AI fuses symbolic and statistical AI. It combines methods from machine	

Recommender About Imprint

#### Try it out: <u>https://rdfox-recommender.poolparty.biz/</u>

#### **Document Splitting PPX APIs** Document/ Page extraction **Data Processing** Triplification Data linking RDF data from Metadata external source are triplification RDF linked to the (activedriectory ID, edit documents date, folder structure) **RDFox database** Storage of RDFox **Data Storage** SPARQL Recommendation

Taxonomy

OntoChem, Mesh

tt

Management

ń

S-Drive

Document

Repositories

#### Questions from User Input PPT APIs Concepts addition PPX APIs through Question extraction autocomplete Fingerprinting Recommendation

#### Architecture

Data Sources



#### Implementation of tf-idf using Datalog rules



<urn:tfidf>(?d, swcs:hasEuclideanNorm, ?euclideanNorm) :-

```
Normalized score
   <urn:data>(?d, ppx:conceptAnnotation, ?x),
                                                                           for document.
AGGREGATE(
   <urn:data>(?d. ppx:conceptAnnotation, ?x),
   <urn:tfidf>(?x, swcs:hasTfIdf, ?tfidf)
   ON ?d
   BIND SUM(POW(?tfidf.2)) as ?squaresSum
),
BIND (SQRT(?squaresSum) AS ?euclideanNorm).
                                                                           tf-idf = tf * idf
<urn:tfidf>(?x, swcs:hasTfIdf, ?tfidf) :-
   <urn:data>(?d, ppx:conceptAnnotation, ?x),
   <urn:data>(?x, ppx:extractedConcept, ?c),
   <urn:data>(?x, ppx:frequency, ?tf),
   <urn:tfidf>(?c, swcs:hasIdf, ?idf),
                                                                           Document frequency
   BIND ((?tf * ?idf) AS ?tfidf).
                                                                            (df) per concept
<urn:tfidf>(?c, swcs:hasIdf, ?idf) :-
AGGREGATE(
   <urn:data>(?d, ppx:conceptAnnotation, ?x),
                                                                           No. of documents.
   <urn:data>(?x. ppx:extractedConcept. ?c)
   ON ?c
   BIND COUNT(distinct ?d) AS ?df
),
AGGREGATE(
                                                                            Inverse document
   <urn:data>(?d, ppx:conceptAnnotation, ?a)
   BIND COUNT(distinct ?d) AS ?n
                                                                           frequency (idf)
),
BIND ((LOG((?n + 1)/(?df + 1)) + 1) AS ?idf).
```

### Part IV: The proof is in the pudding!





© Semantic Web Company

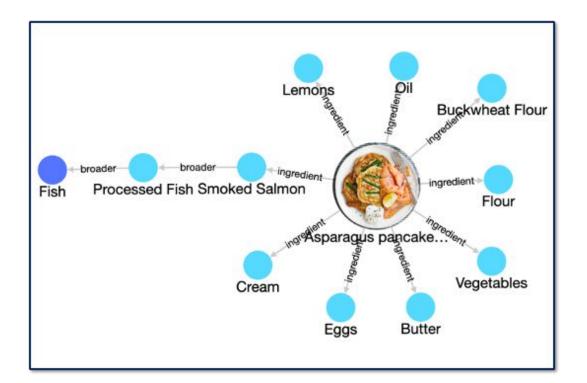
#### Food recommender demo





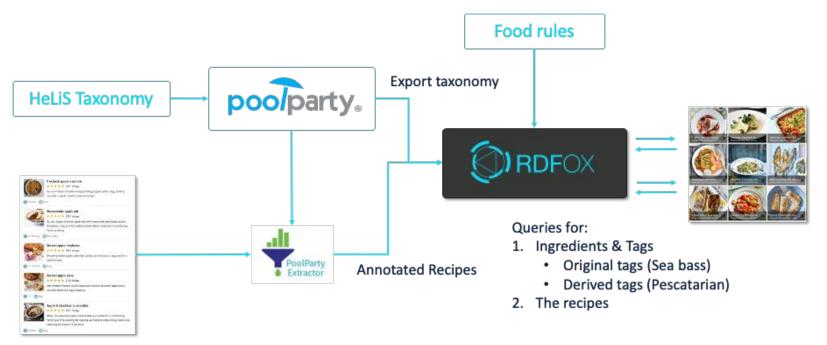
#### **Recipe data model**





#### Architecture

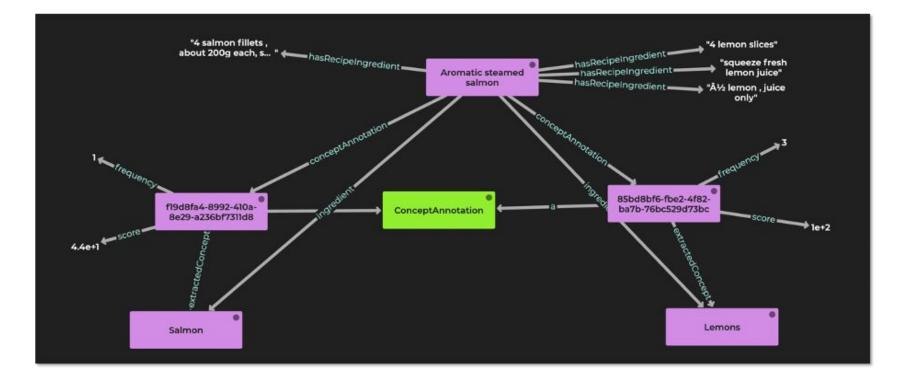




Recipe data: ingredients, tags...

#### **Concept annotation via Extractor**





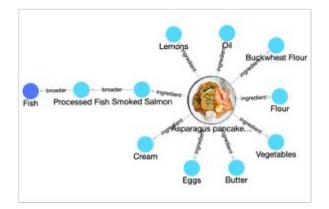
#### **Derived concepts**





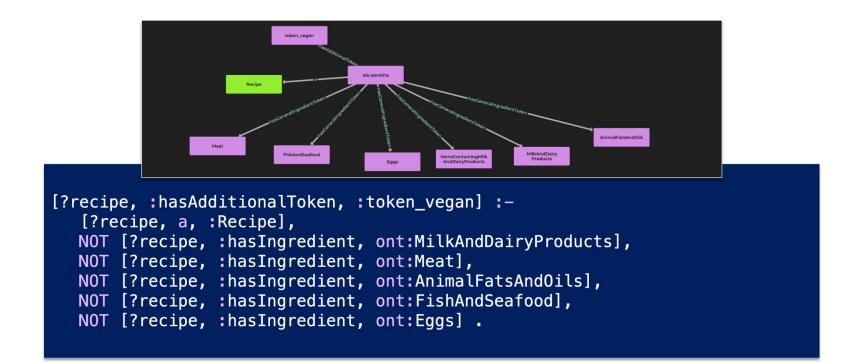
RDFox runs rules to:

- Simplify the annotations graph
  - Adding a direct relationship for ingredient
- · Follow the broader relationships
  - Salmon -> Fresh Fish -> Fish
- Derive new food related concepts
  - Cheese -> Dairy
  - [NOT] Dairy -> Dairy Free



#### Vegan rule using 'negation as failure'





#### Motivation for augmented intelligence



- Can we involve taxonomists/ontologists in adjusting the data for the Datalog rules?
- Can we bring them closer to AI by enhancing decision making with "human-in-the-loop"?

The idea is to make the entire process more user friendly and at the same time be able to inspect and change the results...*this requires rewriting the original Datalog rules into more generic ones*.

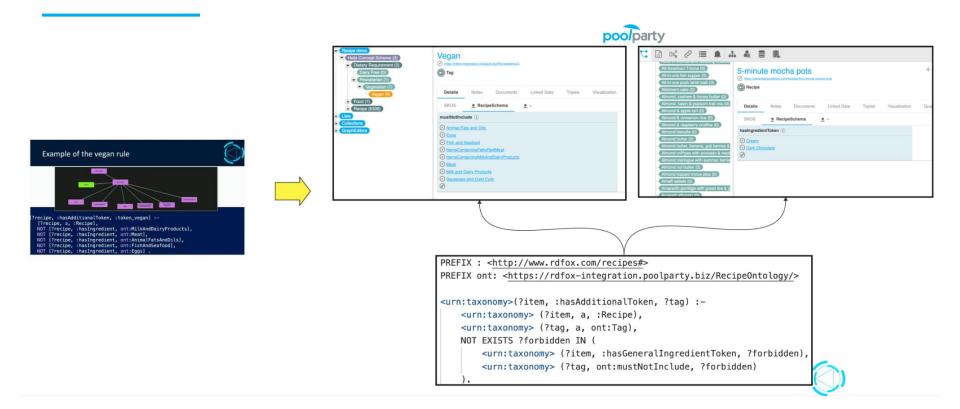


#### • Combining PoolParty and RDFox

- A taxonomist/ontologist is involved directly (human-in-the-loop) by changing the rules and seeing the impact immediately in the application;
- A taxonomist/ontologist would be able to write rules without a necessity to know and understanding the Datalog syntax;
- Managing Datalog rules would be possible from the same interface that PoolParty offers to manage taxonomies and ontologies.

#### Approach





#### Approach (cnt.)

Recipe demo

• Food (7)

Lists

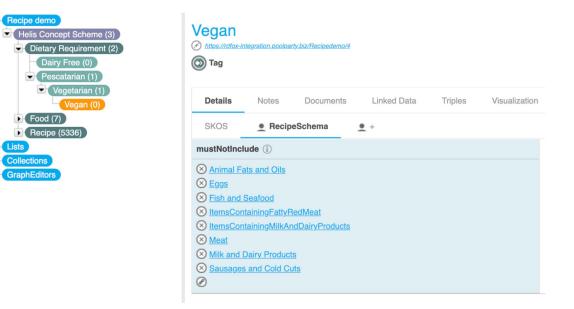
Collections

GraphEditors

• Recipe (5336)

-

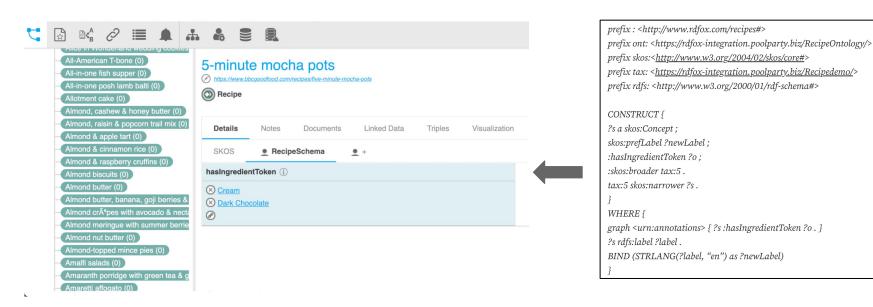
Vegan "Tag" in PoolParty using concepts from Food taxonomy via ont:mustNotInclude relations, specifying what it must not contain.





#### Approach (cnt.)





Recipe "5-minute mocha pots" specifying the ingredients via *:hasIngredientToken* relationships by using concepts from the Food taxonomy.



• The new rewritten Datalog rule enables knowledge engineers to accommodate their input:

PREFIX : <http://www.rdfox.com/recipes#>
PREFIX ont: <https://rdfox-integration.poolparty.biz/RecipeOntology/>

<urn:taxonomy>(?item, :hasAdditionalToken, ?tag) :-<urn:taxonomy> (?item, a, :Recipe), <urn:taxonomy> (?tag, a, ont:Tag), NOT EXISTS ?forbidden IN ( <urn:taxonomy> (?item, :hasGeneralIngredientToken, ?forbidden), <urn:taxonomy> (?tag, ont:mustNotInclude, ?forbidden) ).

### Approach (cnt.)



#### • We also take broaders into account to preserve correctness:

<urn:taxonomy>(?recipe, :hasGeneralIngredientToken, ?token) :-<urn:taxonomy>(?recipe, :hasIngredientToken, ?token) .

<urn:taxonomy>(?recipe, :hasGeneralIngredientToken, ?broaderToken) :-<urn:taxonomy>(?recipe, :hasGeneralIngredientToken, ?token), <urn:taxonomy>(?token, skos:broader, ?broaderToken).



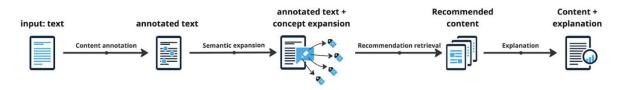


- Use explain command to verify why a fact has been derived
- A courtesy of *symbolic-based approach* to AI.

Syntax:

explain [shortest] [<max\_depth> [<max\_rule\_inst>]] <fact>

Description: This command explains how a fact has been derived.



Demo



