

Linked Data for Digital Twins

Jan Top, Hannelore Heuer, December 14, 2022



Digital twin

- Digital model of something, for example a human being
- Can tell us something about this person that we do not see immediately
- Can relate to scientific knowledge
- Is used for forecasting, visualisation, diagnosis, advice, control



Data for digital twins

■ For creating the twin

- Learn from response of the real system to external stimuli
- Learning can happen in advance or in real time
- Example data: food intake, triglyceride and sugar levels, ...
- Twins also use expert knowledge!

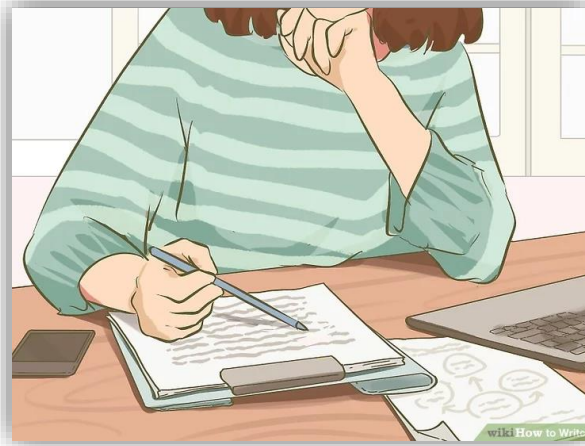
■ For feeding the twin

- Adapt twin model or parameter values in the model
- Monitor current status of the subject and its environment
- Example data: heart beat, blood pressure, weight, ...
- The twin cannot (and should not try to) predict everything

Digital twin for nutritional advice

What could a digital twin of your field do for you?

Which data or knowledge would it need?



Data sharing and reuse – a complex process

- Internal and external resources
- Legacy systems
- Data quality (fit-for-use) never perfect
 - Different objectives
 - Different versions
 - ...

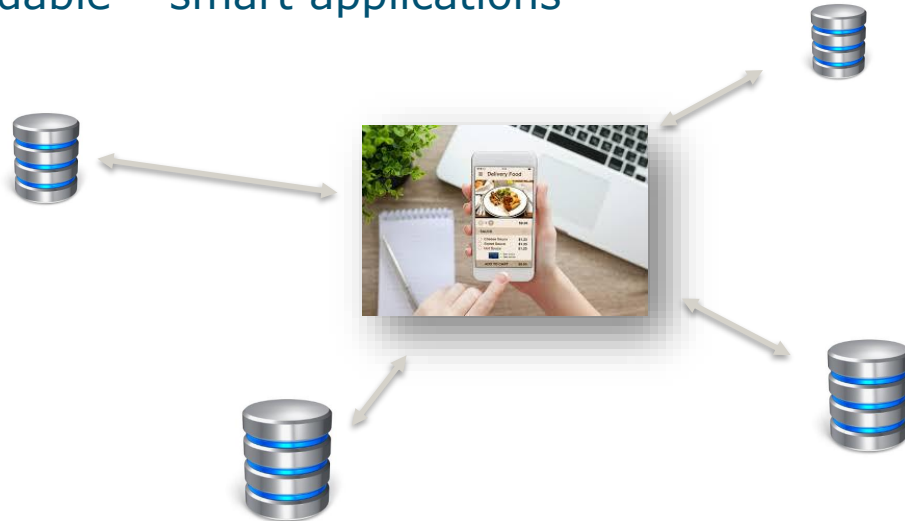


Manual selection and combination: laborious, costly and error prone

The ideal picture



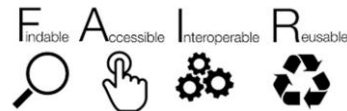
- Direct access to primary sources
- Get it right at the source (smart devices)
- Access to what is needed for a specific task and user
- Machine readable – smart applications



How to share information?



- Know that a resource exists and where to find it
- Being able to access it and read it
- Ensure that machines can read it such that they show correct behaviour
- Ensure that it can be used in multiple contexts



The FAIR principles

- Enable sharing information on the web
- Raw data, processed data, algorithms, software, documents, ...
- FAIR is not equal to 'open'
- FAIR can contribute to 'fair'

But also: link to other datasets



Example: linking food product data



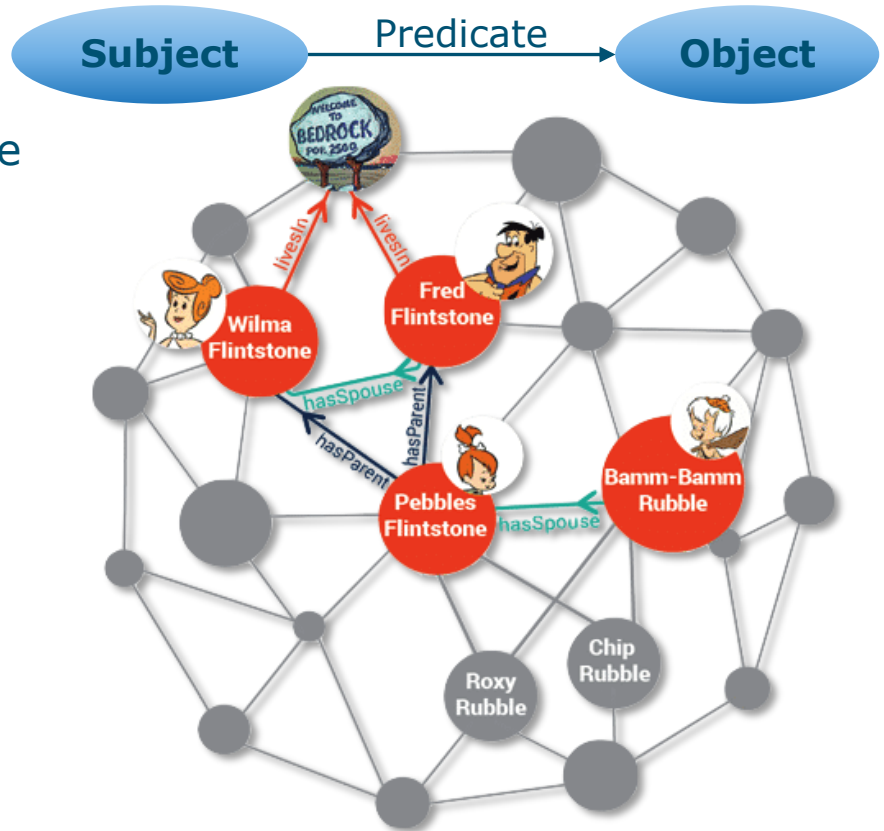
	C	D	K	L	P	R	S	T	U	V
1	Productcode	Product_omschrijving	Meeteen	HoeveelH	ENERCJ_kj	PROT_g	PROTPL_g	PROTAN_g	NT_g	CHO_g
2	1	Aardappelen rauw	g	100	371	2	2	0		19
3	2	Aardappelen nieuwe rauw	g	100	371	2	2	0		19
4	3	Aardappelen oude rauw	g	100	371	2	2	0		19
5	121	Aardappelpuree vers bereid m hv melk m mar	g	100	350	2.2	1.4	0.9		13.9
6	667	Cassave rauw	g	100	657	0.6	0.6	0	0.1	36.8
7	668	Taro rauw	g	100	505	1.4	1.4	0	0.2	26.2
8	669	Yam rauw	g	100	526	1.5	1.5	0	0.3	28.2
9	670	Pomtajer rauw	g	100	577	2	2	0		31
10	671	Aardappel zoete rauw	g	100	413	1.2	1.2	0	0.2	21.3
11	787	Aardappelpuree instant- gem bereid	g	100	249	1.4	0.7	0.8		12.5
12	856	Rosti onbereid	g	100	518	2.5	2.5	0		21.5
13	948	Rosti bereid z vet	g	100	626	1.9	1.9	0		22.8
14	982	Aardappelen z schil gekookt gem	g	100	353	1.9	1.9	0		17.4

	A	B	K	M	N	O	P	C
1	Food_code	Product_brand	Preparation_method	no_sweet	m_sweet	sd_sweet	se_sweet	no_sc
2	121	AH aardappelpuree met boter	microwave 5 minutes 850 watt	10	7	6	2	
3	948	Aviko Rösti rondjes naturel	oven 220 degrees, 15 minutes	9	9	7	2	
4	982	AH krieltjes	boil for 10 minutes	12	5	3	1	
5	1150	AH Aardappelschijfjes	fry with 15% Blue Band Iedere dag voor koken, bal	7	6	6	2	
6	1456	Aviko Pommes Frites	deep-fry in Diamant original frituurvet at 175 °C fo	9	9			
7	2325	AH Vastkokende aardappelen	boil for 20 minutes	7	6			
8	2834	AH aardappelbolletjes	deep-fry in Diamant original frituurvet at 175 °C fo	10	7			
9	383	Appelsientje Goudappel		12	46			
10	395	Coca cola regular		12	56	10	3	
11	400	Seven up regular		12	51	12	3	
12	410	Appelsientje sinaasappelsap		12	28	10	3	
13	411	Spa blauw		11	1	1	0	
14	413	AH Tomatensap		9	10	8	3	
15	414	Royal Club Tonic		8	24	8	3	



Linking data points as triples

- Data are measurements, observations, facts, ..., that can be structured as **triples**
- Entities and properties require **global identifiers**
- Numbers, strings remain anonymous



Play-a-LOD

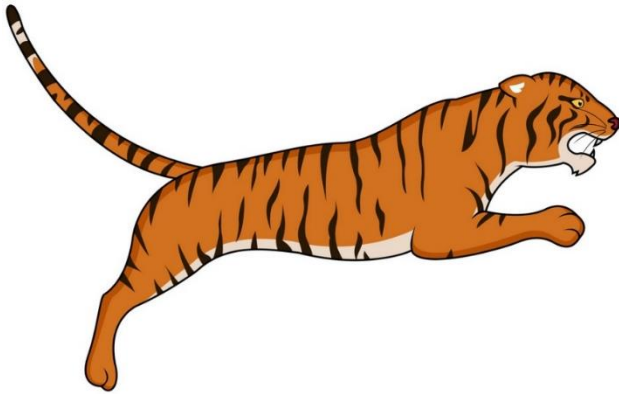


Play-a-LOD



Play-a-LOD

- Take nine cards each from the stack
- Build as many triples as you can, extending the current graph
- Take new cards, until nine

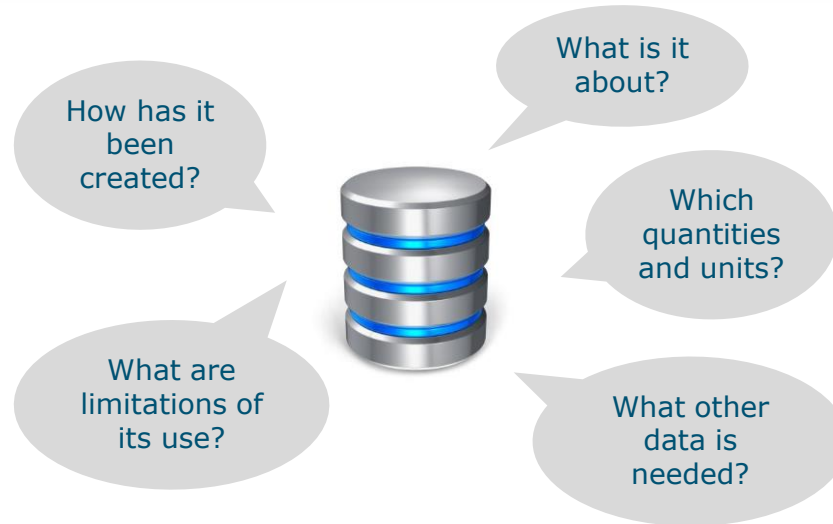


is a

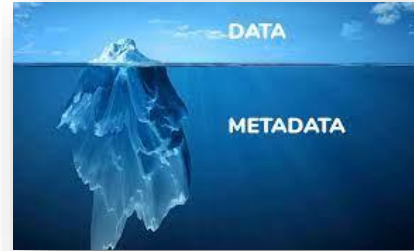


Data without meaning is meaningless

01001	BUTTER, WITH SALT	15.87	717	0.85	81.11	2.11	0.06	0	0.06
01002	BUTTER, WHIPPED, WITH SALT	15.87	717	0.85	81.11	2.11	0.06	0	0.06
01003	BUTTER OIL, ANHYDROUS	0.24	876	0.28	99.48	0	0	0	0
01004	CHEESE, BLUE	42.41	353	21.4	28.74	5.11	2.34	0	0.5
01005	CHEESE, BRICK	41.11	371	23.24	29.68	3.18	2.79	0	0.51
01006	CHEESE, BRIE	48.42	334	20.75	27.68	2.7	0.45	0	0.45
01007	CHEESE, CAMEMBERT	51.8	300	19.8	24.26	3.68	0.46	0	0.46
01008	CHEESE, CARAWAY	39.28	376	25.18	29.2	3.28	3.06	0	
01009	CHEESE, CHEDDAR	37.1	406	24.04	33.82	3.71	1.33	0	0.28



Metadata



- **Explanation:** for proper use
 - Observed properties, quantities, units
 - Which objects, events or materials are being observed
 - External conditions, circumstances

- **Provenance:** for reliability, reproduction, tracing
 - 'Library' metadata: creator, date, institute, ...
 - Acquisition method, experiment, production method, ...

Metadata – machine readable

Editor About

Repository Catalog Dataset Distribution

Show optional fields

Title

License

Has version

Access URL

Download URL

Media type

Part of dataset

Build Share

Select a field to read more about it.

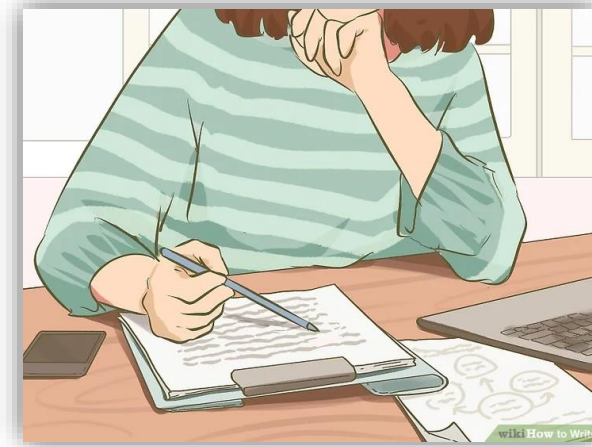
RDF preview

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix dct: <http://purl.org/dc/terms/>.
@prefix dcat: <http://www.w3.org/ns/dcat#>.
@prefix fdp: <http://rdf.biosemantics.org/ontologies/fdp-0#>.
@prefix datacite: <http://purl.org/spar/datacite/>.

<> rdf:type dcat:Distribution;
    dct:title "Food Nutritional Values";
    dct:license <http://purl.org/NET/rdflicense/cc-by1.0>;
    dct:hasVersion "1.2";
    dcat:accessURL <http://wur.nl>;
    dcat:mediaType "Excel".
```

Your metadata

How do you currently describe your data?



Metadata: use a shared vocabulary or ontology

<http://www.foodvoc.org/resource/FoodTaxonomy/BlueCheese>

<http://www.ontology-of-units-of-measure.org/resource/om-2/Mass>

	A	B	C	D	E	F	G	H	I	J
1	NDB_No	Shrt_Desc	Water_(g)	Energ_Kcal	Protein_(g)	Lipid_Tot_(g)	Ash_(g)	Carbohydrt_(g)	Fiber_TD_(g)	Sugar_Tot_(g)
2	01001	BUTTER,WITH SALT	15.87	717	0.85	81.11	2.11	0.06	0	0.06
3	01002	BUTTER,WHIPPED,WITH SALT	15.87	717	0.85	81.11	2.11	0.06	0	0.06
4	01003	BUTTER OIL,ANHYDROUS	0.24	876	0.28	99.48	0	0	0	0
5	01004	CHEESE,BLUE	42.41	353	21.4	28.74	5.11	2.34	0	0.5
6	01005	CHEESE,BRICK	41.11	371	23.24	29.68	3.18	2.79	0	0.51
7	01006	CHEESE,BRIE	48.42	334	20.75	27.68	2.7	0.45	0	0.45
8	01007	CHEESE,CAMEMBERT	51.8	300	19.8	24.26	3.68	0.46	0	0.46
9	01008	CHEESE,CARAWAY	39.28	376	25.18	29.2	3.28	3.06	0	
10	01009	CHEESE,CHEDDAR	37.1	406	24.04	33.82	3.71	1.33	0	0.28

Controlled vocabulary (typically for search)

ROC+

Users Remote Sources Admin Info

Select Project Add Terms Identify Synonyms Create Taxonomy

Add Terms

Add term Add

Page size: 20

Term	Synonyms
ACHILLES-HEEL	
acibenzolar-S-methyl	Acibenzolar-S-methyl (fr), Acibenzolar-s-metile (it)
acid derivatives	derivados ácidos (es), derivati acidi (it), dérivés d'acides (fr)
acid rain	happosade (fi), kwaśna deszcz (pl), lluvia ácida (es), piogge (nl)
acid soil	acidic soil (en), gleba kwaśna (pl), hapan maaperä (fi), low pH soil (it), zure grond (nl), suelo ácido, suolo acido, zure g
acid solution treatment	traitement en solution acide (fr), tratamiento con solución ácido agente acidificante (it), trattamenti con soluzioni acide (it), trattamento con soluzione acida (it), zuuroplo
acidic additive to manure	aanzuringsproduct (nl), additif acidifiant aux engrais organiques (fr), dodatek do nawozu (pl), Zuur toevoeging aan mest (nl)
acidification	acidificación (es), acidification (fr), acidificazione (it), happan
ACOSIC	
acoustic sensor	akustinen sensori (fi), capteur acoustique (fr), geluidsensoren (nl), acustico (it)

acid soil

Synonyms : acidic soil, gleba kwaśna, hapan maaperä, low pH soil, sol acide, suelo ácido, suolo acido, zure g

Narrower relation : --

Broader relation : soil type by acidity

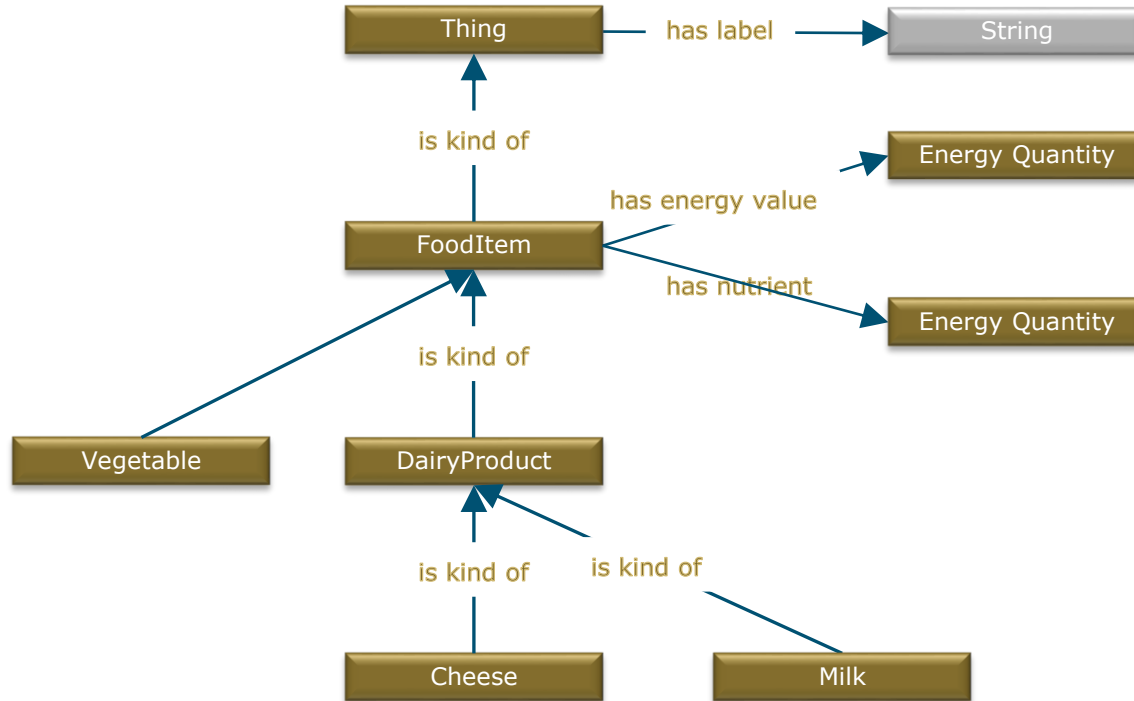
Related to : --

Edited on 2016-10-06

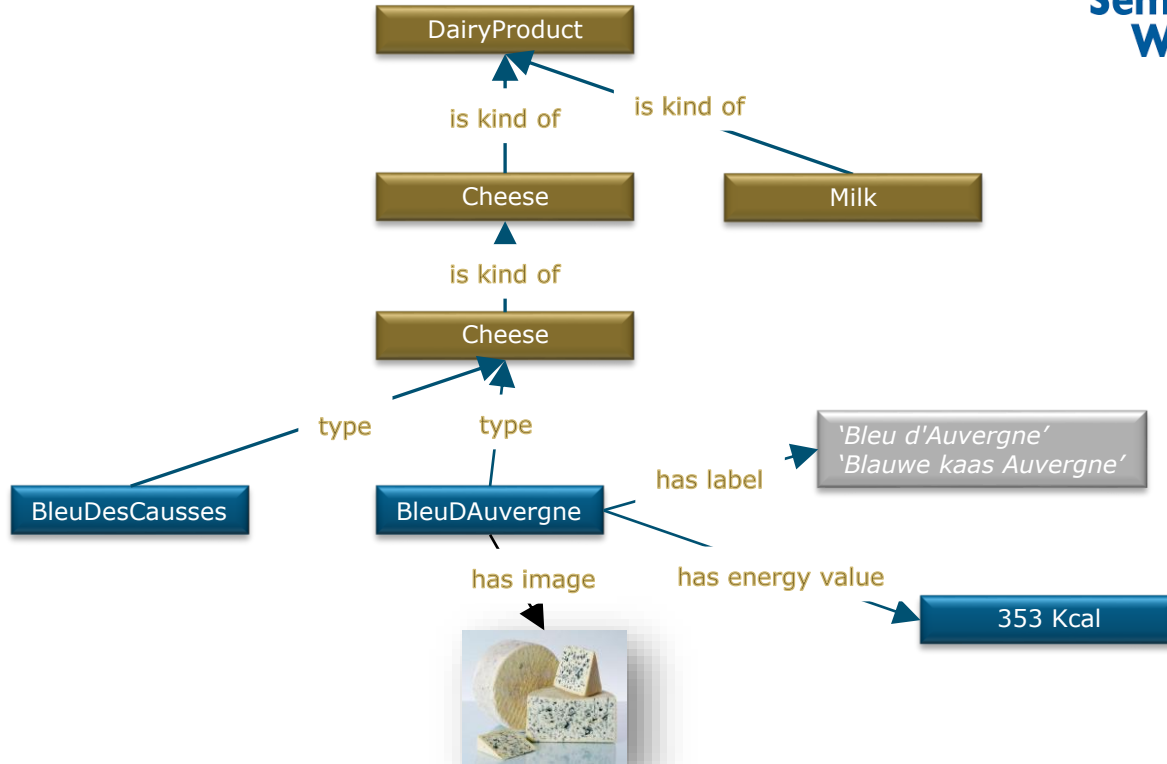
- biogeochemical cycle
- biogeographic region
- biological interaction
- biotope
- cause of climate change
- chemical efficacy
- chemical element
- climate
 - arid climate
 - climate by Köppen
 - continental subarctic climates with extremely severe winters
 - continental subarctic or boreal (taiga) climates
 - desert climate
 - dry-summer maritime subalpine climate
 - hot summer continental climates
 - ice cap climate
 - maritime subarctic climates or subpolar oceanic climate
 - Mediterranean climate



Ontology



Knowledge graph: ontology and data



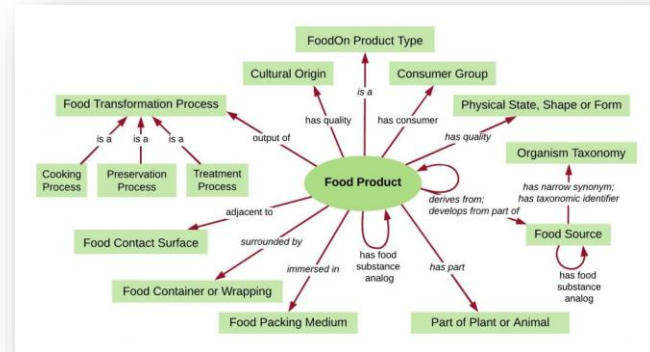
Current ontologies and vocabularies

- [OBO Foundry](#)
- [Ontobee](#)
- [The Ontology Lookup Service](#)
- [Bioportal](#)
- [Agroportal](#)

AGROVOC Linked Open Data

Linked Data is a method of web publication in which each individual piece of data is:

- uniquely identified using [HTTP URIs](#) (that is, URLs, or 'web addresses'),
- available both as 'machine readable' data and as 'human readable pages, and
- linked to other resources.

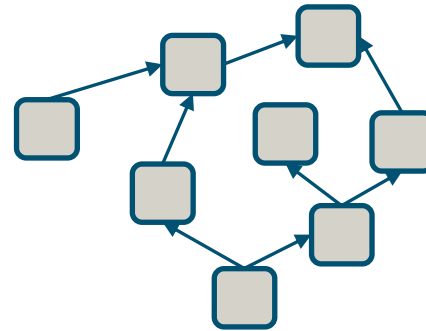
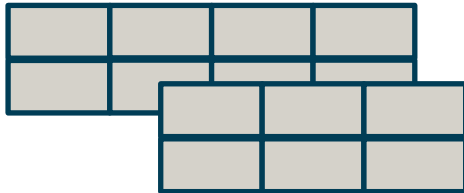


Linked data versus relational data

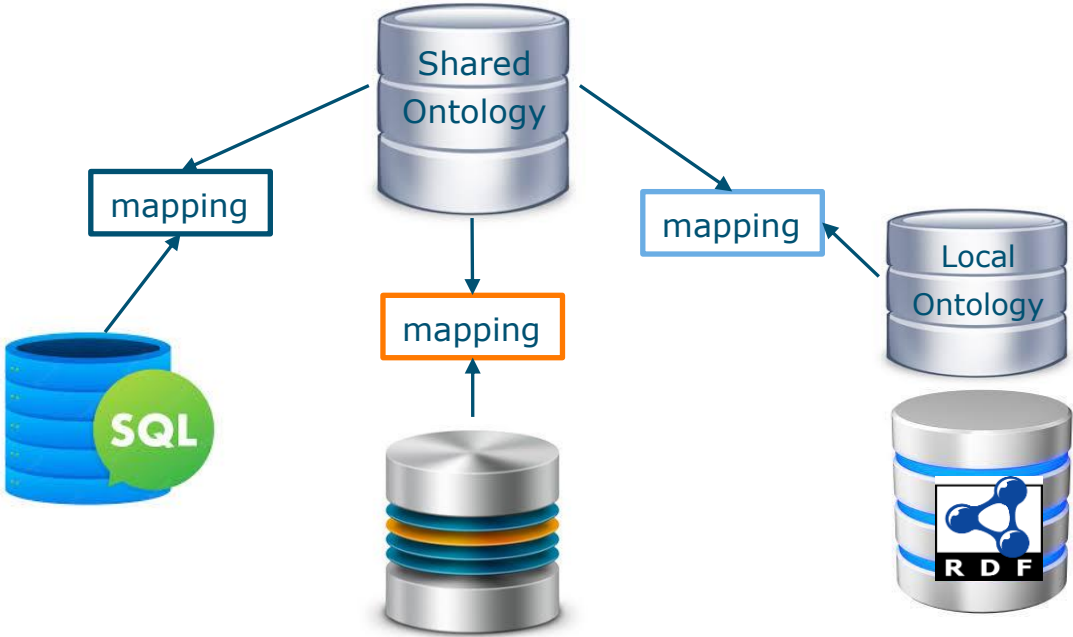
No pre-set table structure, so easier to add just any other fact; no need to update the database design

Table headers are data as well, allowing automatic merging and reasoning without human intervention

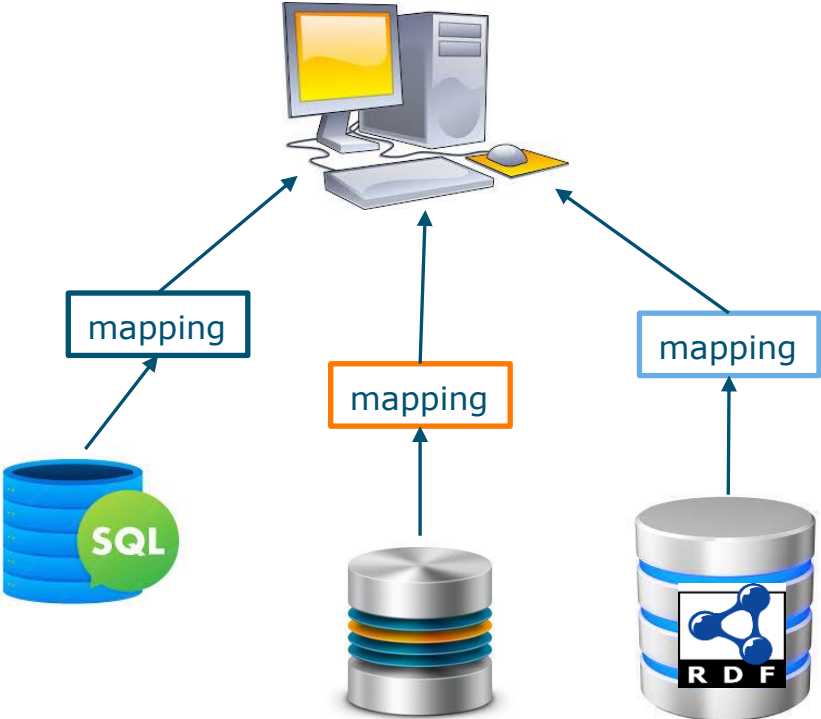
Combining relational data with semantic wrapper: best of both worlds



Merging data sources



Merging data sources



It's happening already

schema.org

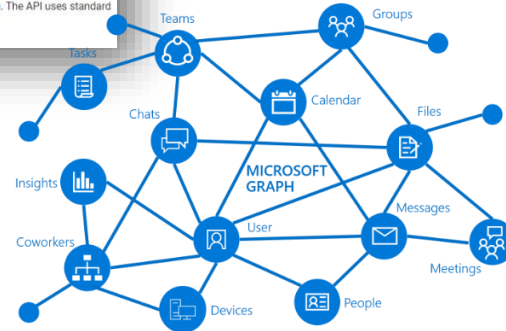
Welcome to Schema.org

Schema.org is a collaborative, community activity with a mission to create, r data on the Internet, on web pages, in email messages, and beyond.

Schema.org vocabulary can be used with many different encodings, includin vocabularies cover entities, relationships between entities and actions, and documented extension model. Over 10 million sites use Schema.org to marl applications from Google, Microsoft, Pinterest, Yandex and others already u experiences.

Graph Database Market Worth \$2.9 Billion by 2024 - Exclusive Report by MarketsandMarkets™

CISION PR Newswire September 16, 2019



Digital twins need linked data and semantics

