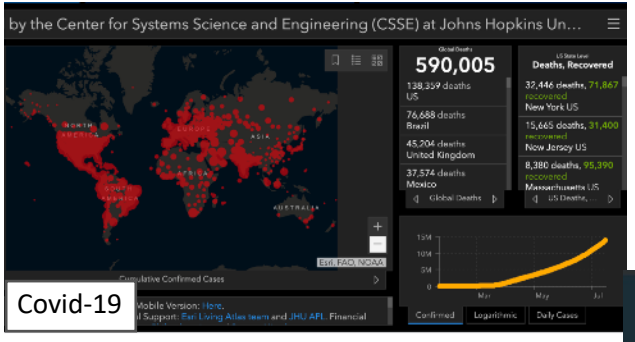
Geospatial Data Integration and Analysis using Virtual Knowledge Graphs

Linfang Ding
KRDB Research Centre for Knowledge and Data
Free University of Bozen-Bolzano, Italy
ding@inf.unibz.it

KRDB Summer Online Seminars
12 June 2020 – Bolzano, Italy

unibz

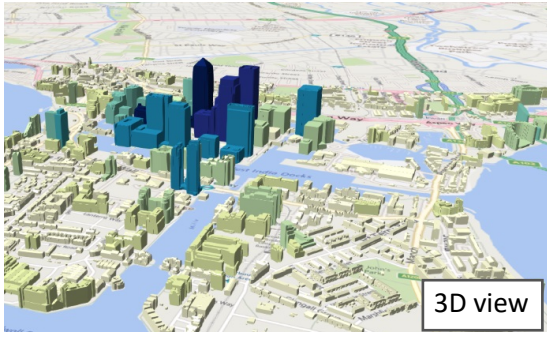


Covid-19

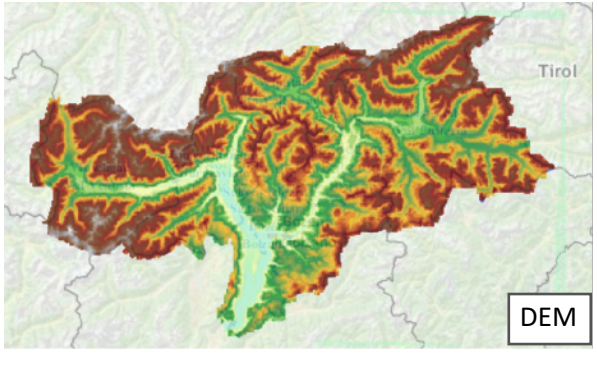
2D map

Geospatial data
is everywhere.

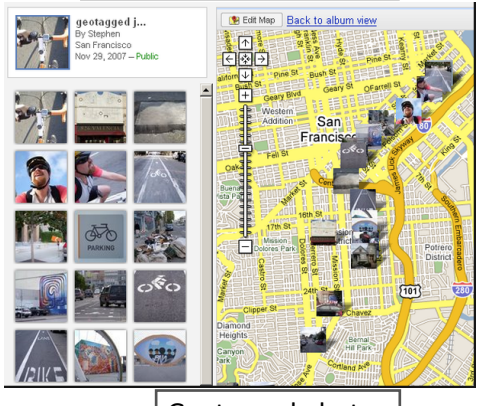
Oceanary transportation



3D view



DEM

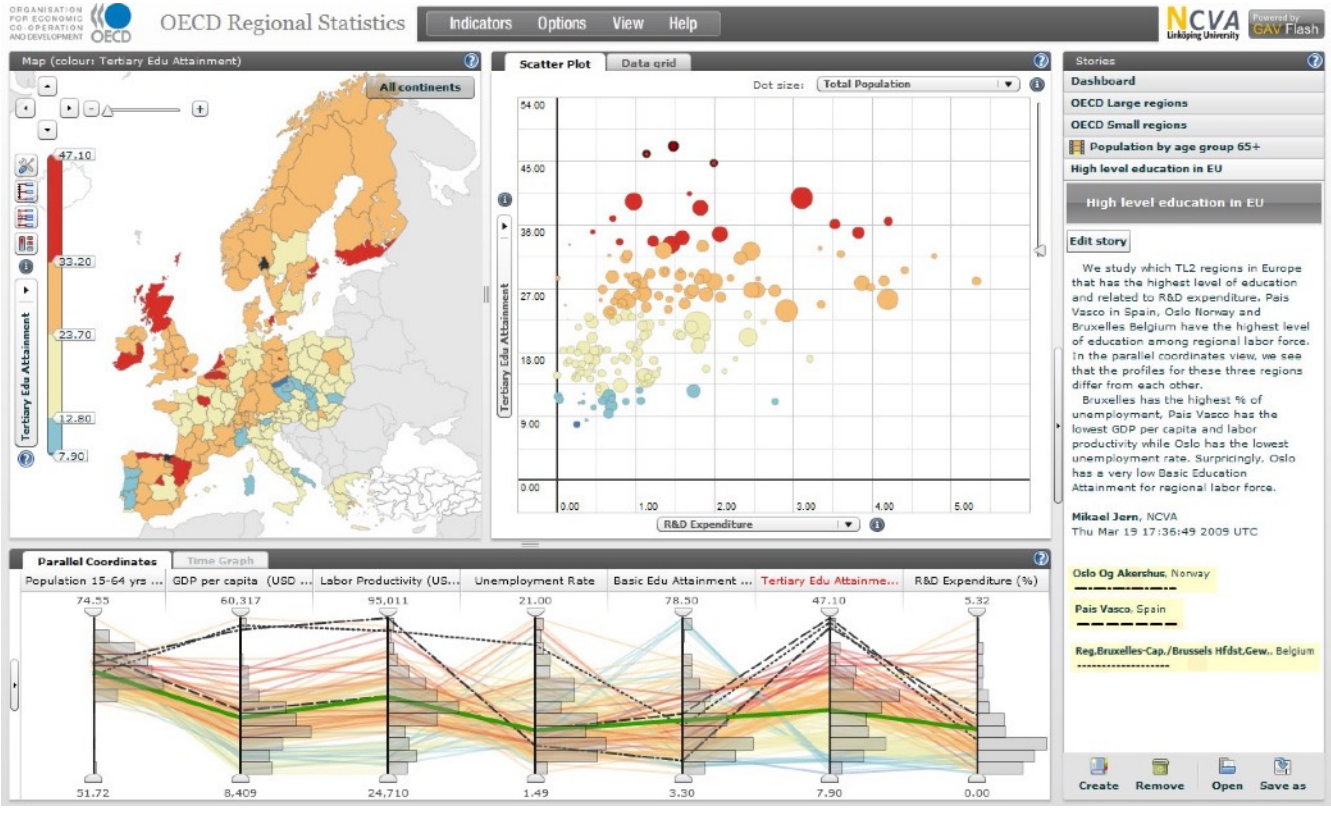


Geotagged photos

unibz

Geovisual Analytics

Interactive graphic exploration of large and complex spatial data facilitated by automated computational methods.



https://stats.oecd.org/OECDregionalstatistics/#story=0

unibz

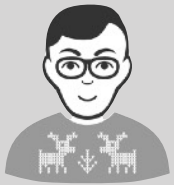
Outline

- Challenges in Geovisual analytics
- Framework
- Use cases
- Conclusion

domain

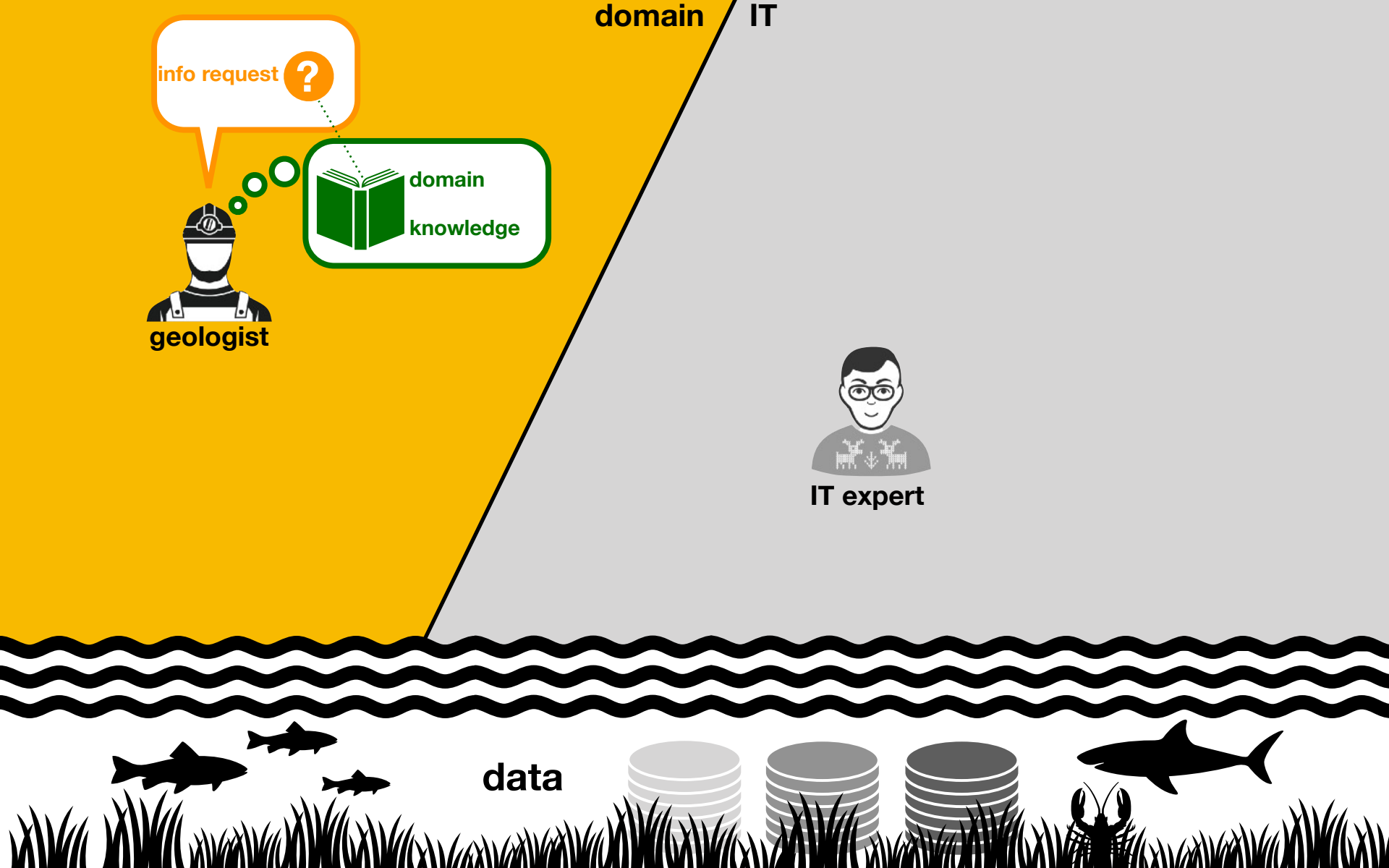
IT

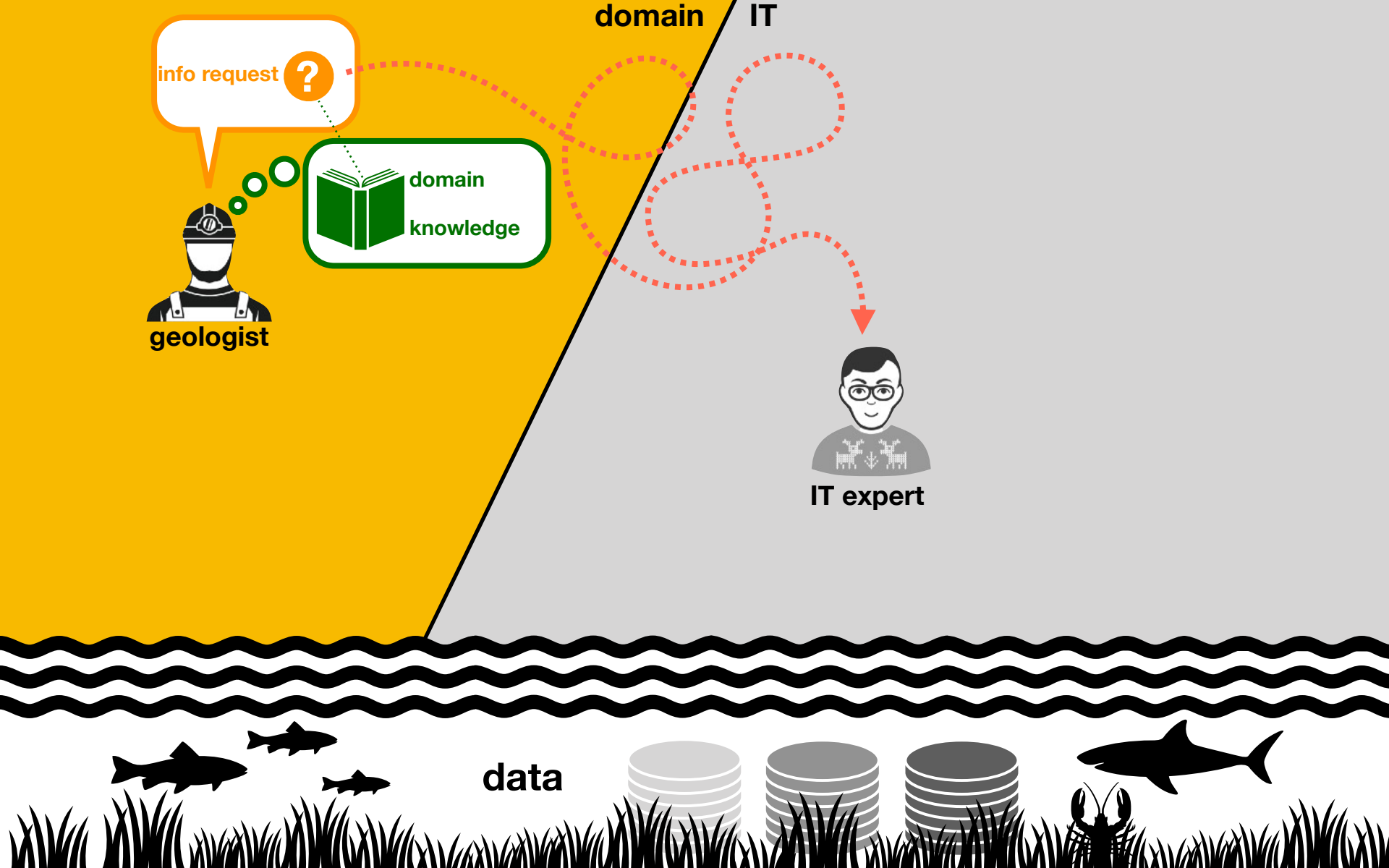
geologist

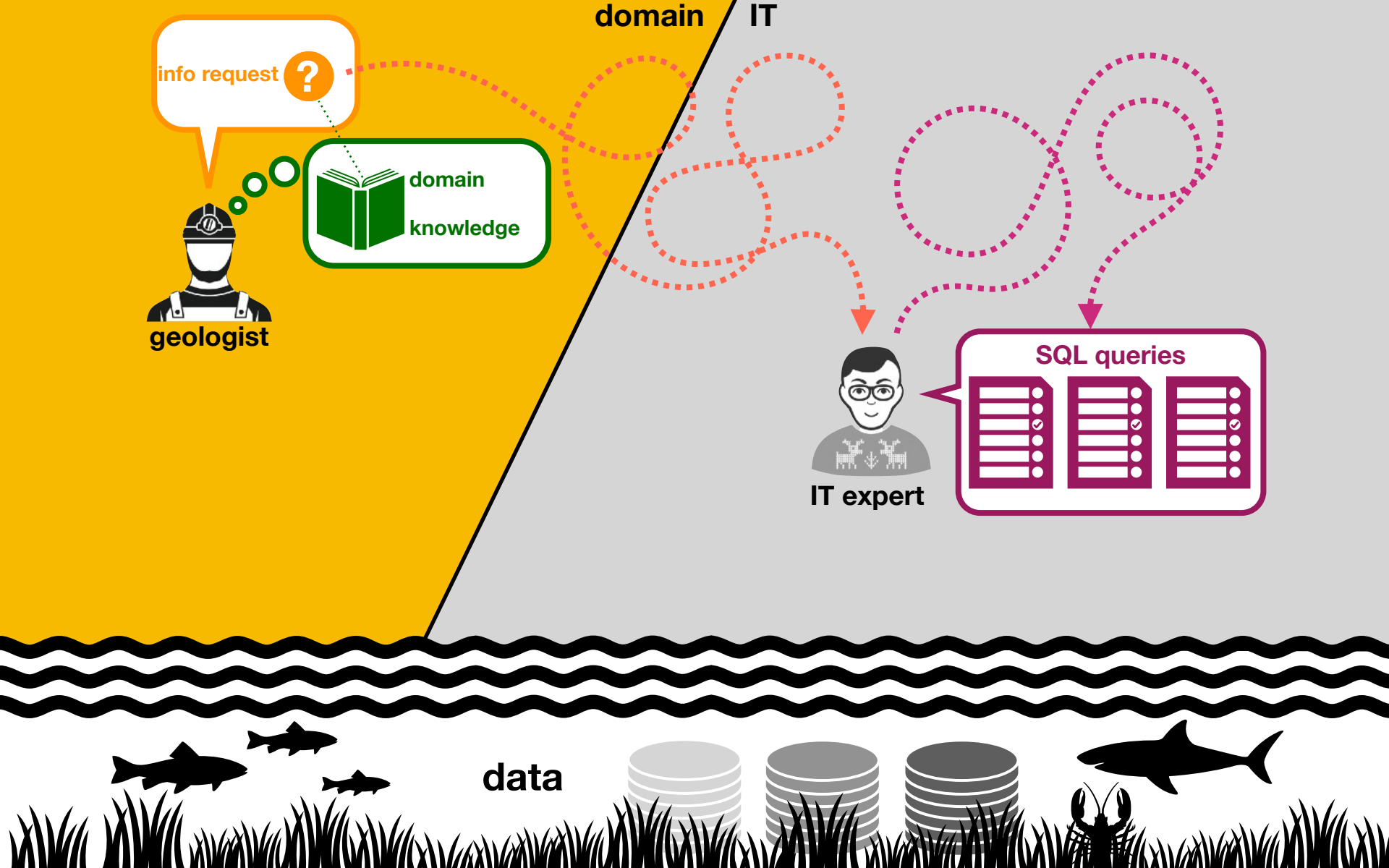


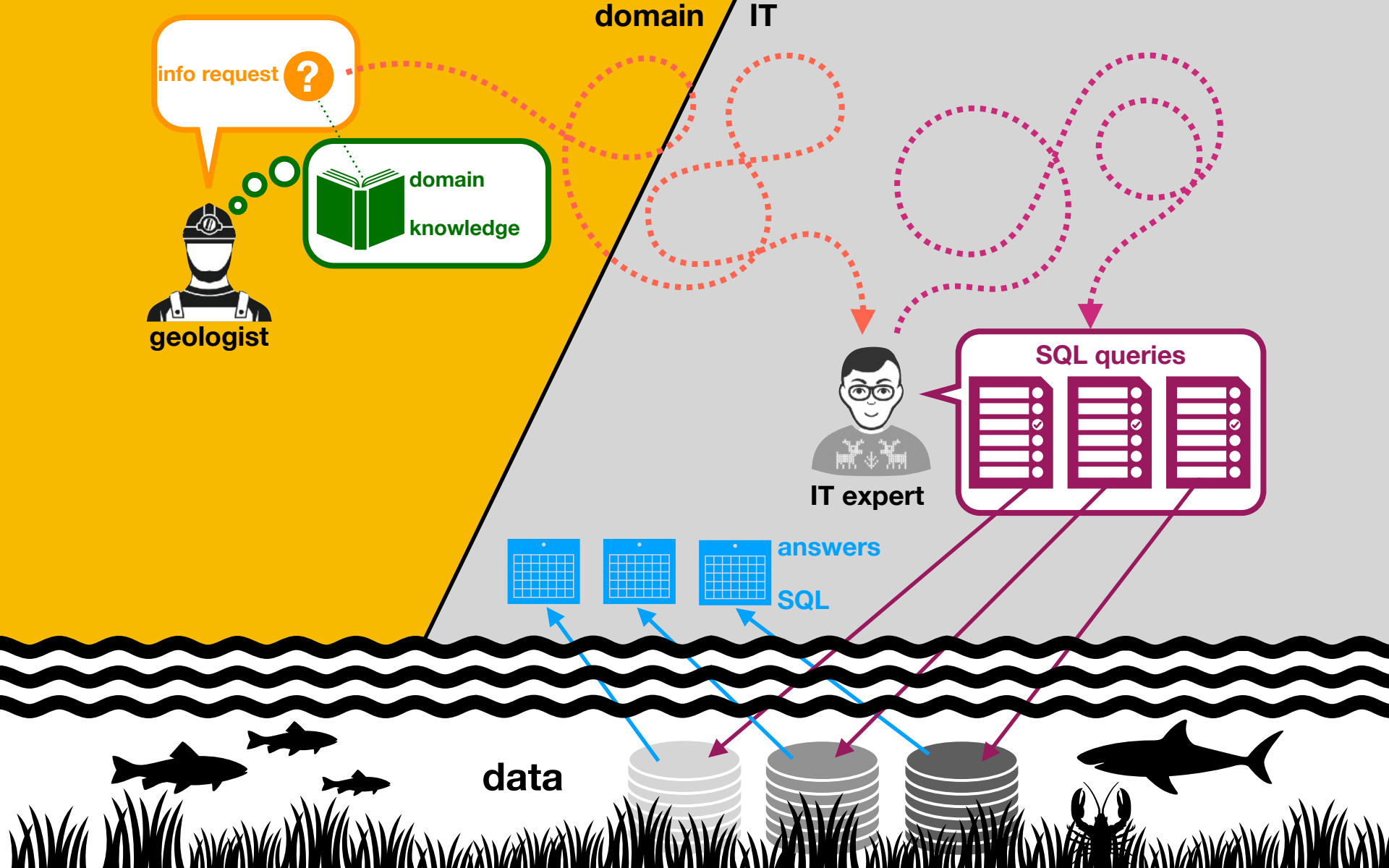
IT expert

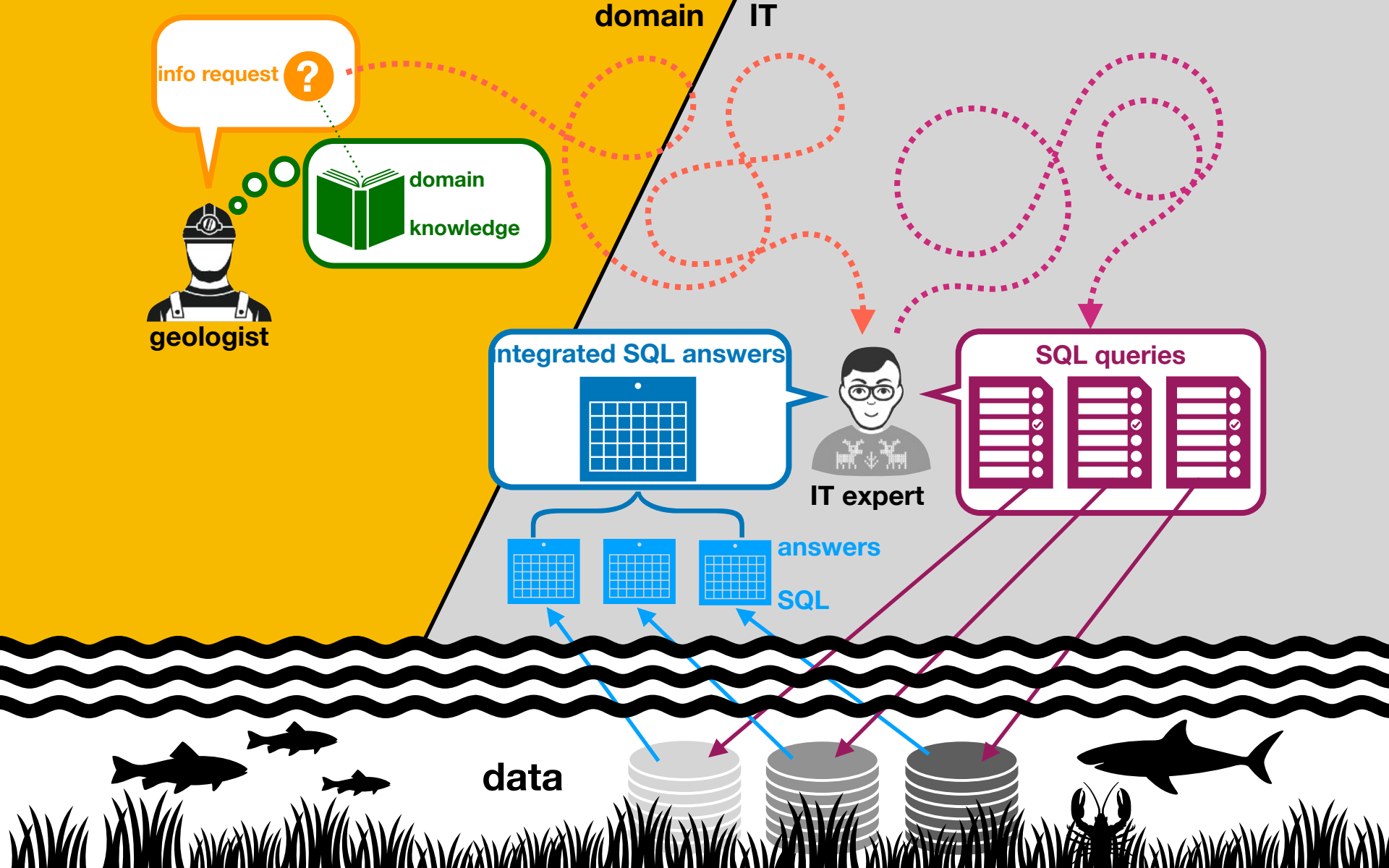
data

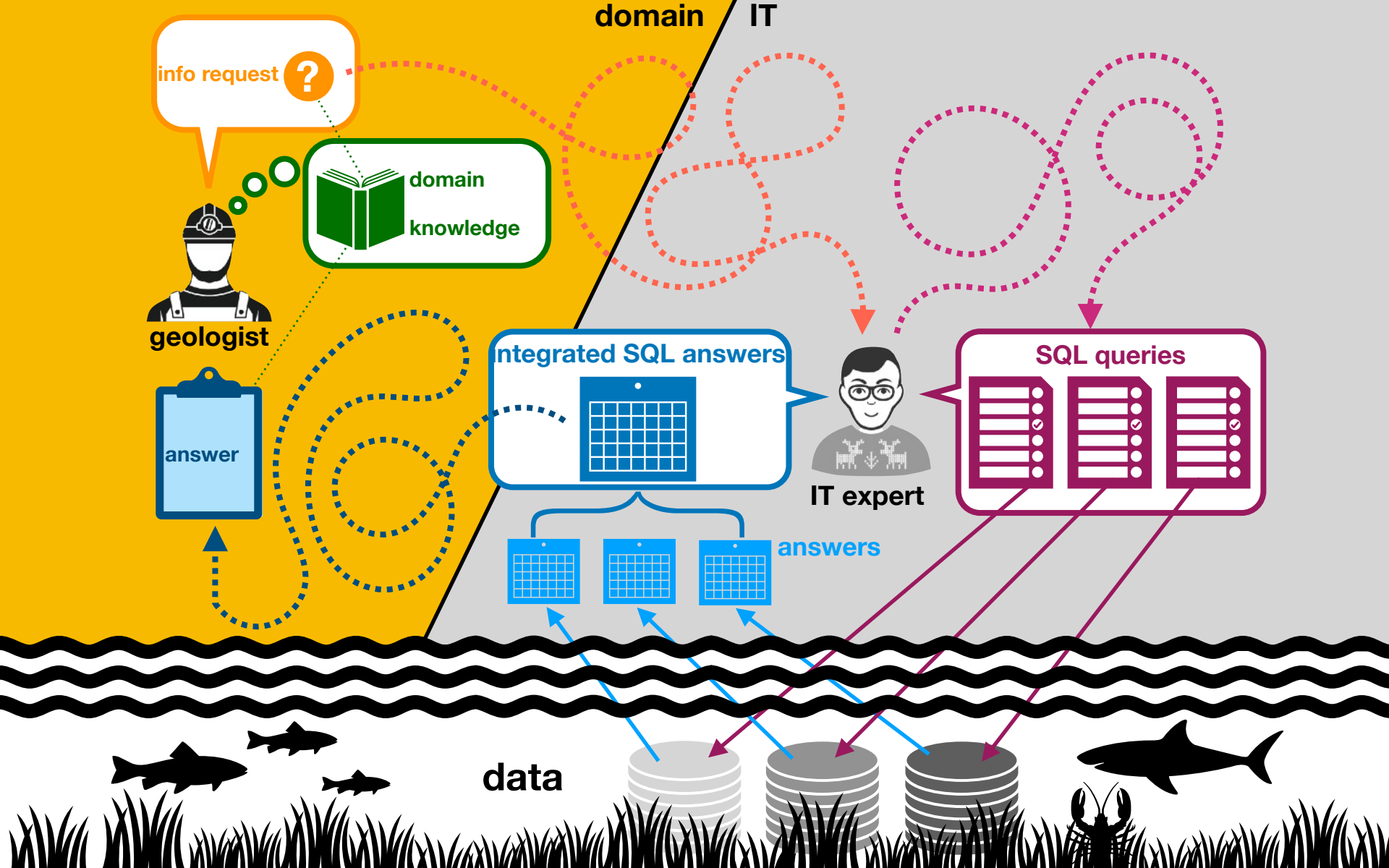

domain
IT
info request
?
domain
knowledge
geologist
IT expert
data

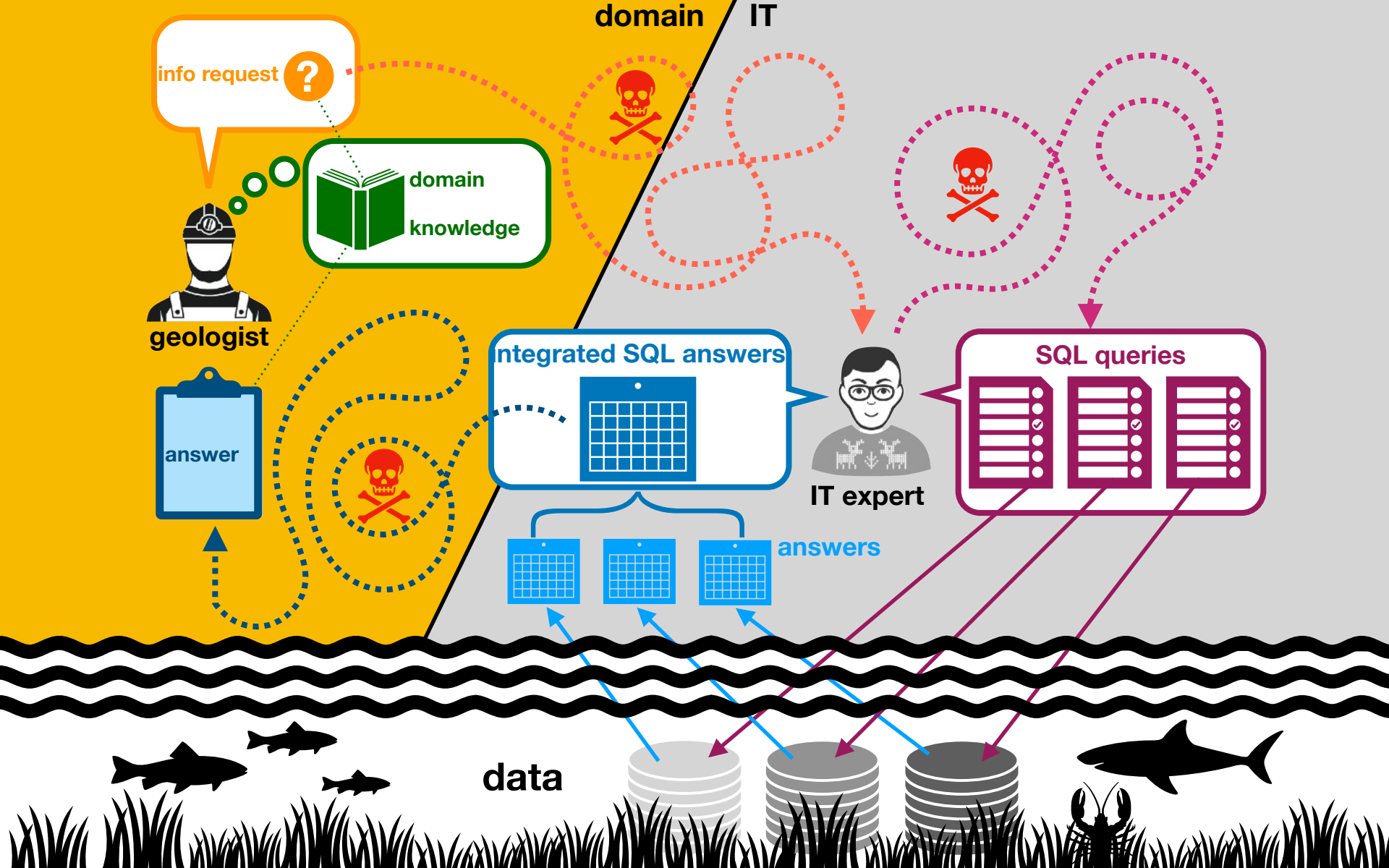

domain
IT
info request
?
domain
knowledge
geologist
IT expert
data

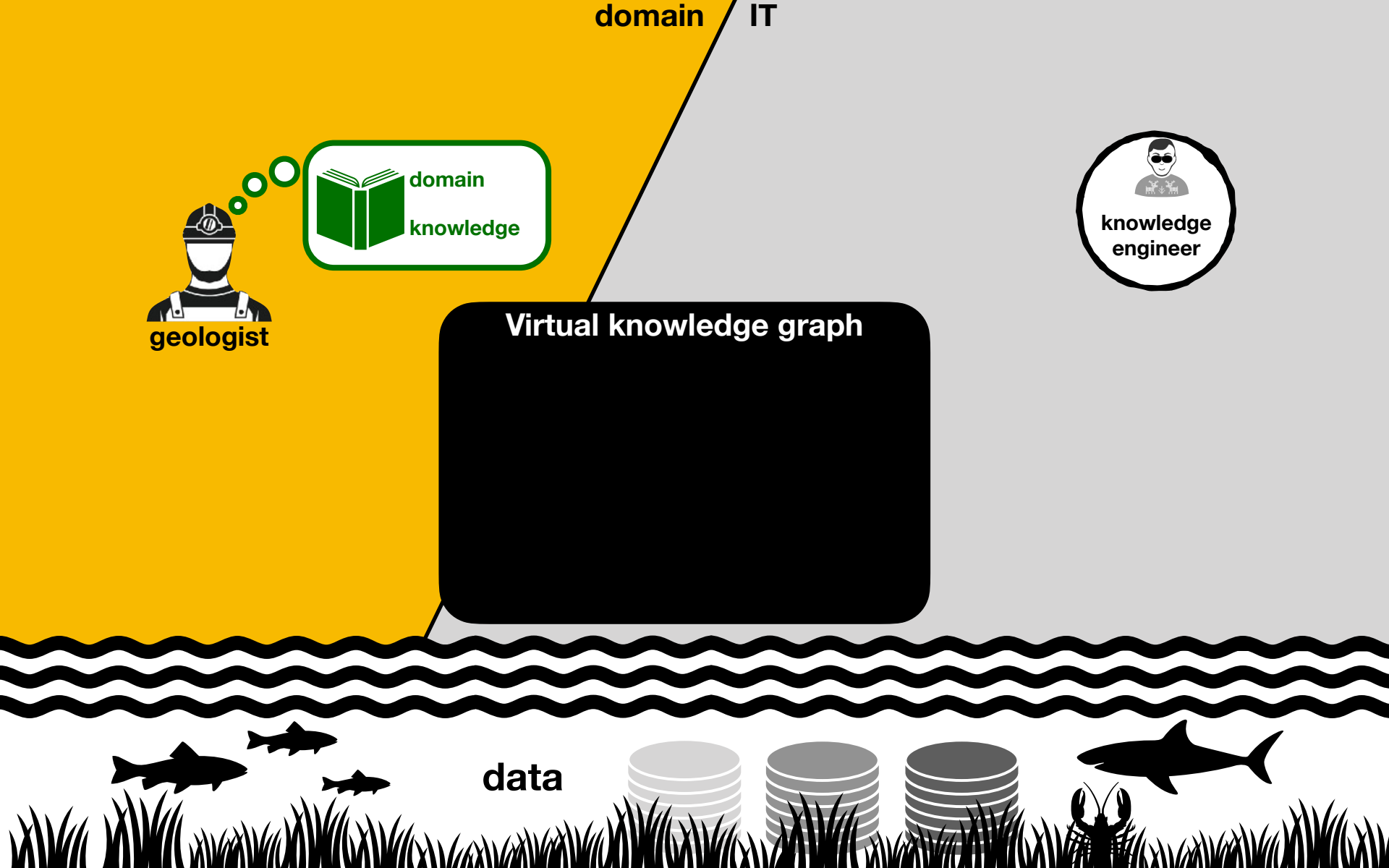

domain
IT
info request
?
domain
knowledge
geologist
IT expert
SQL queries
data

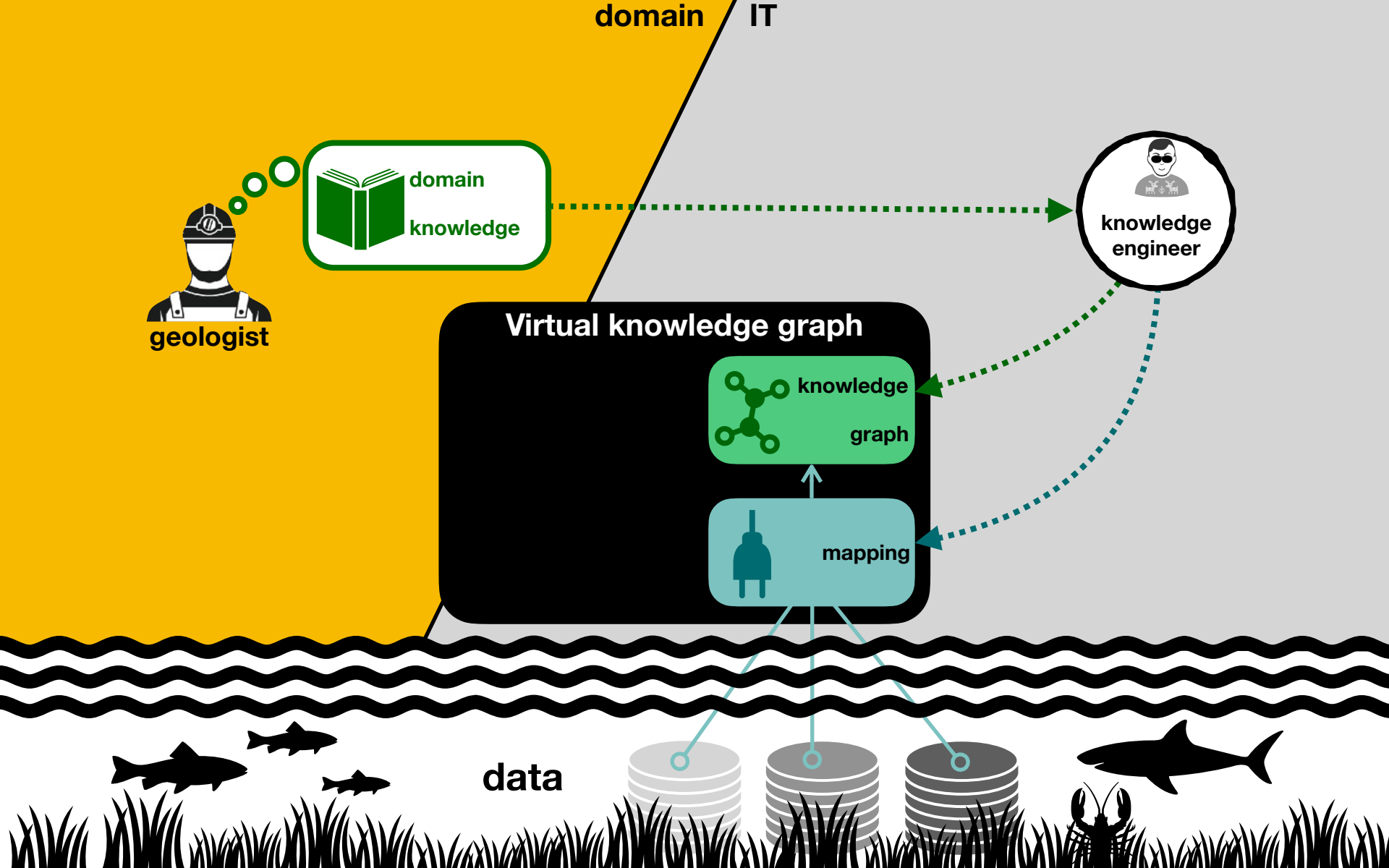

domain
IT
info request
?
domain
knowledge
geologist
IT expert
SQL queries
answers
SQL
data

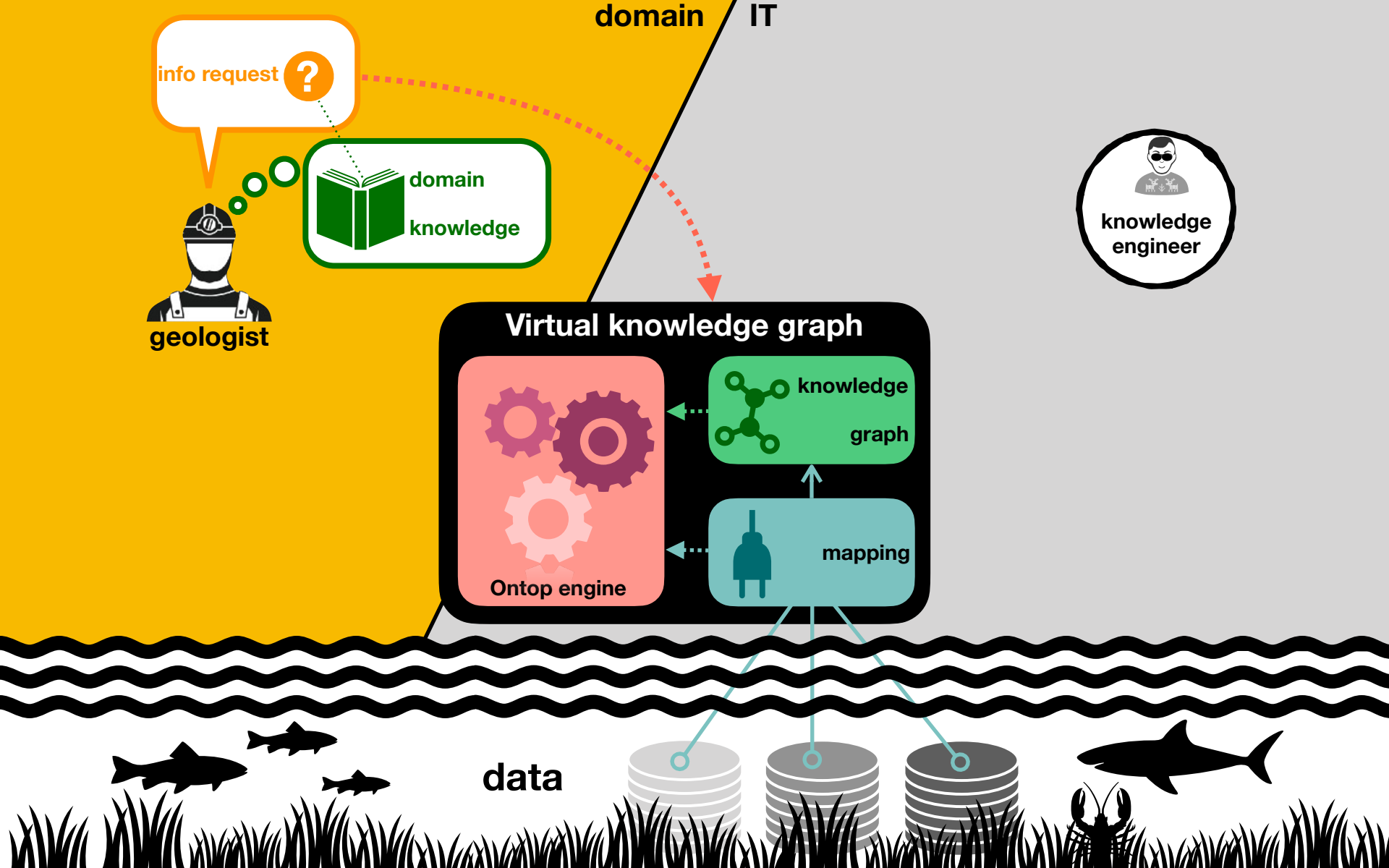

domain
IT
info request
?
domain
knowledge
geologist
integrated SQL answers
IT expert
SQL queries
answers
SQL
data

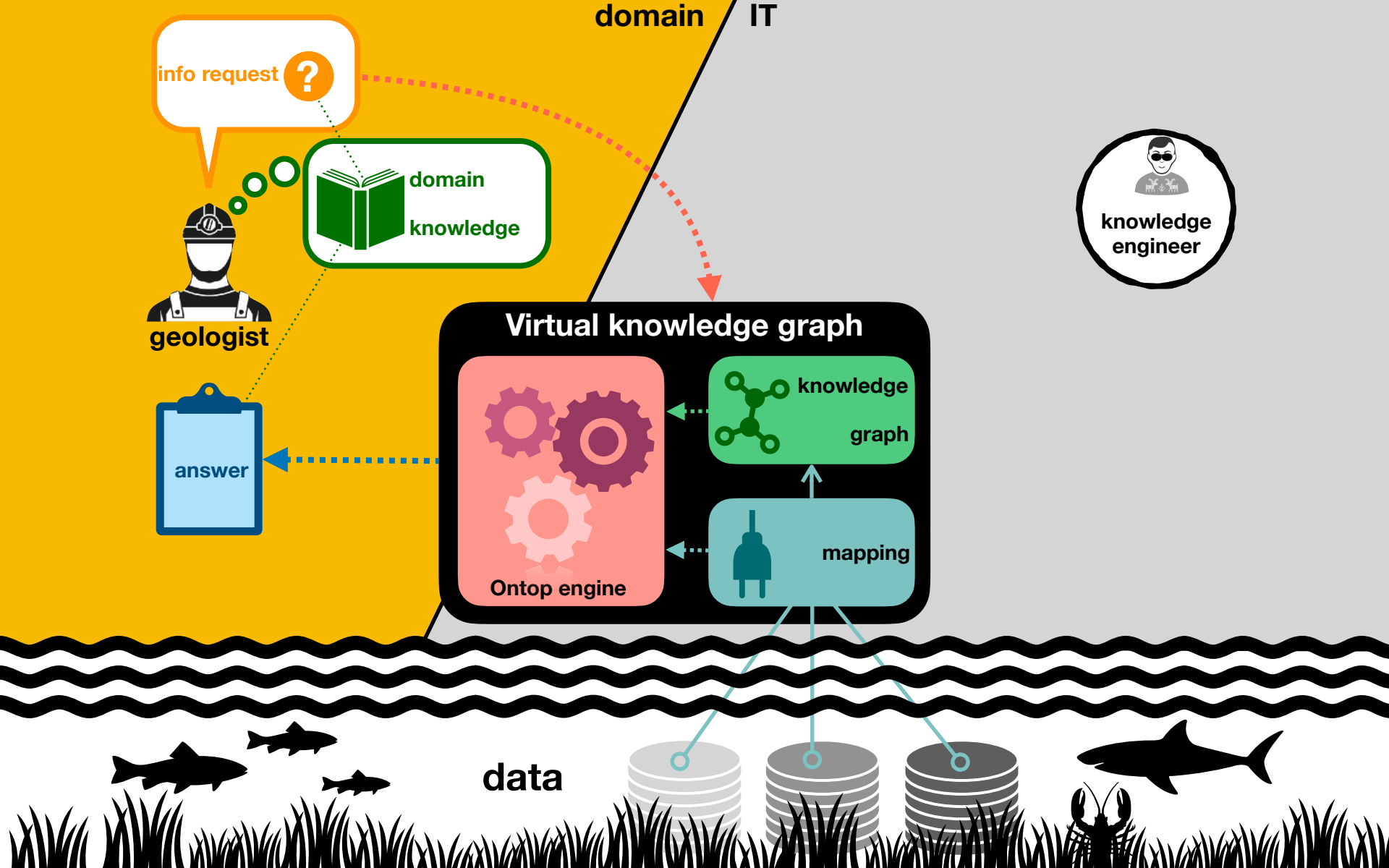

domain
IT
info request
?
domain
knowledge
geologist
answer
integrated SQL answers
IT expert
SQL queries
answers
data


domain
IT
info request
?
domain
knowledge
geologist
answer
integrated SQL answers
IT expert
SQL queries
answers
data


domain
IT
domain
knowledge
geologist
knowledge
engineer
Virtual knowledge graph
data


domain
IT
domain
knowledge
geologist
knowledge
engineer
Virtual knowledge graph
knowledge
graph
mapping
data


domain
IT
info request
?
domain
knowledge
geologist
knowledge
engineer
Virtual knowledge graph
knowledge
graph
mapping
Ontop engine
data

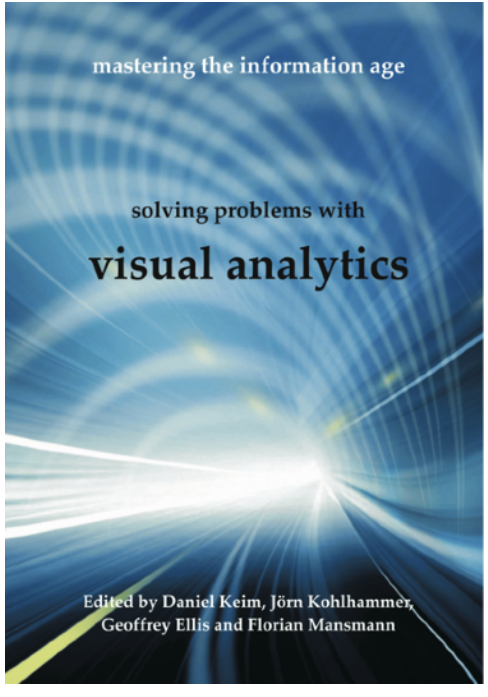

domain
IT
info request
?
domain
knowledge
geologist
knowledge
engineer
Virtual knowledge graph
knowledge
graph
mapping
Ontop engine
answer
data

Challenges in (Geo)Visual analytics

- Effective visual analysis requires large and complex data organized in a coherent way and with clear semantics.

  "*logic based systems, balancing expressive systems power and computational cost represent state of the art solutions. Visual analytics can greatly benefit from such an approach*"

  "*associated with data integration activities, is the need for managing all the data semantics in a centralised way, for example, by adding a virtual logic layer on top of the data itself*".



Keim, D.A.; Kohlhammer, J.; Ellis, G.; Mansmann, F., Eds. *Mastering the Information Age - Solving Problems with Visual Analytics*; Eurographics, 2010.

unibz

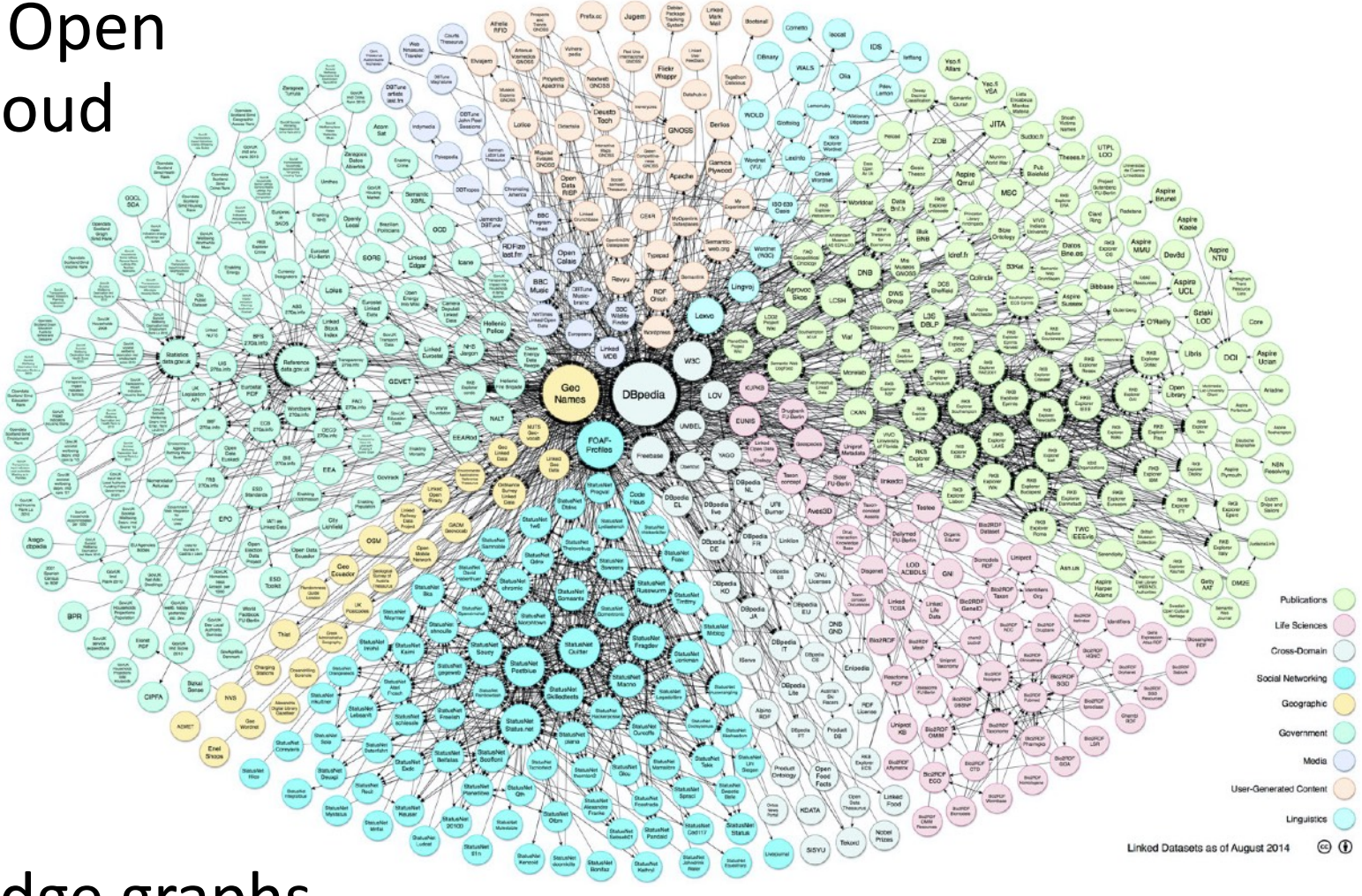
Challenges in Geovisual analytics

"*the significant analytical potential that can come from diverse data representing different perspectives on a problem*"

"while the integration of geospatial big data is a problem, location can be used as a common denominator, and the linked data concept is also promising".

Robinson, A.C.; Demšar, U.; Moore, A.B.; Buckley, A.; Jiang, B.; Field, K.; Kraak, M.J.; Camboim, S.P.; Sluter, C.R. Geospatial big data and cartography: research challenges and opportunities for making maps that matter. *Int. J. of Cartography* **2017**, *3*, 32–60. doi:10.1080/23729333.2016.1278151.

Linked Open Data cloud

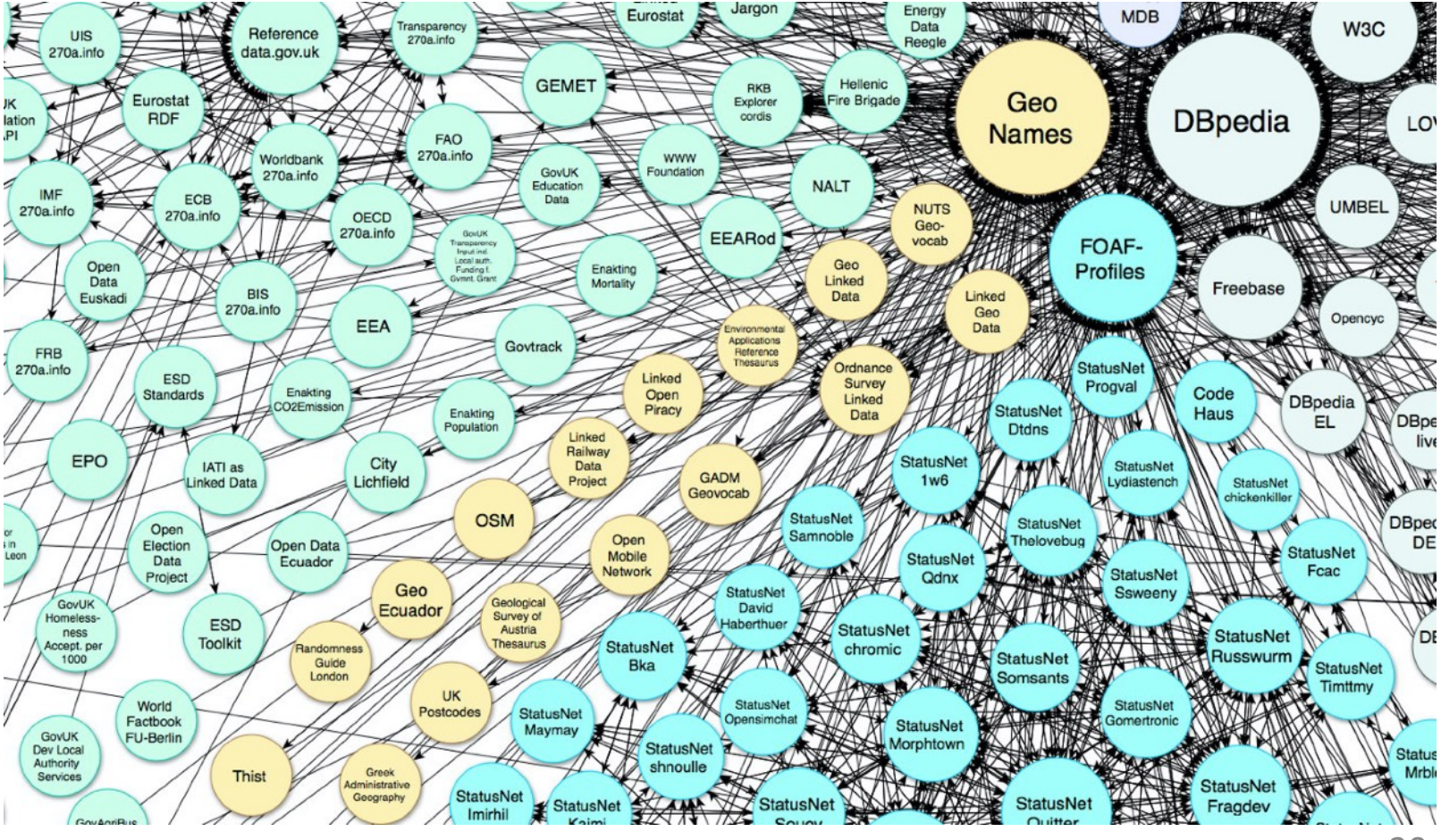


Knowledge graphs

August 30, 2014 (http://lod-cloud.net/)

unibz

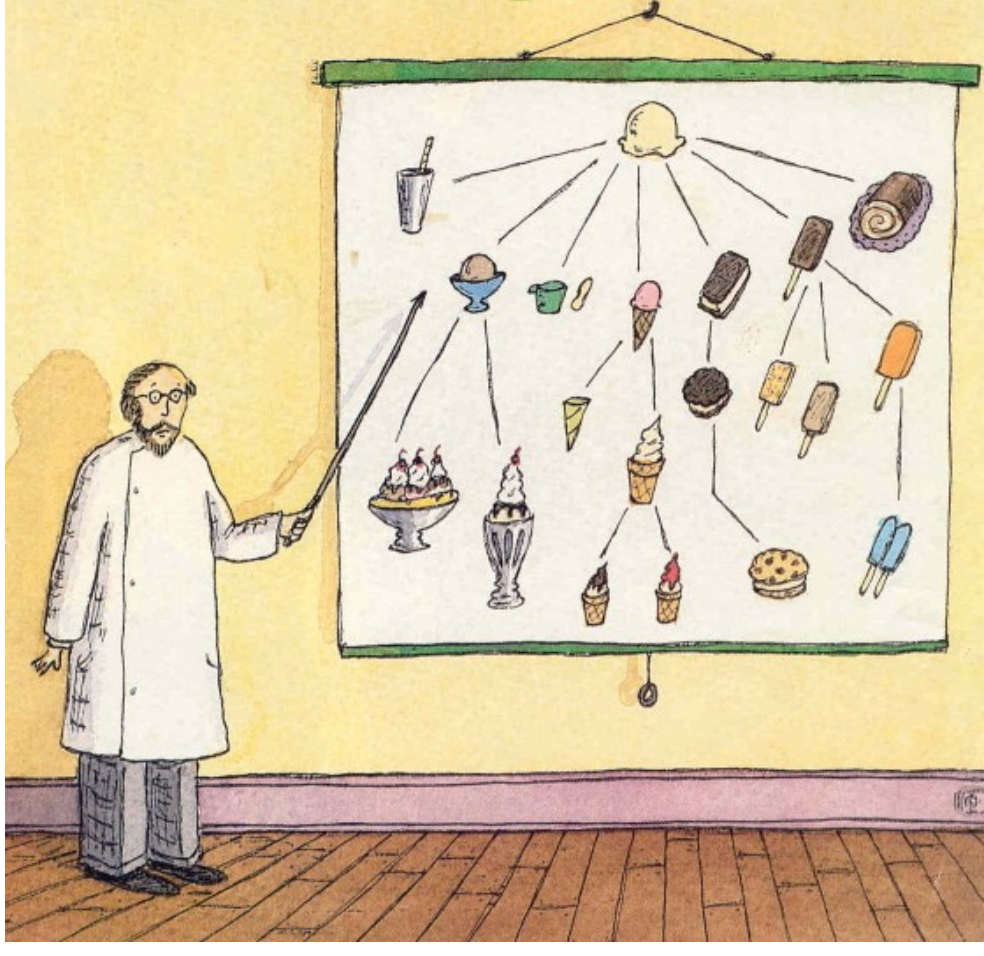
Linked Open Data cloud – GeoSpatial enabled



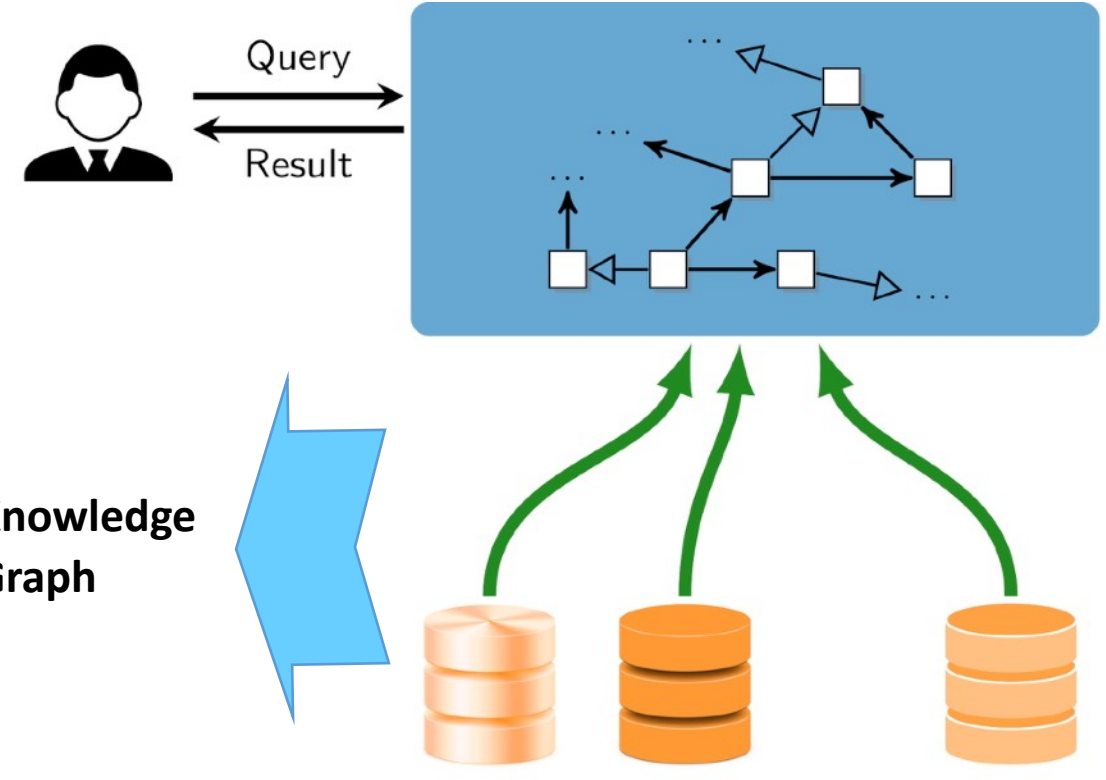
unibz

Ontology

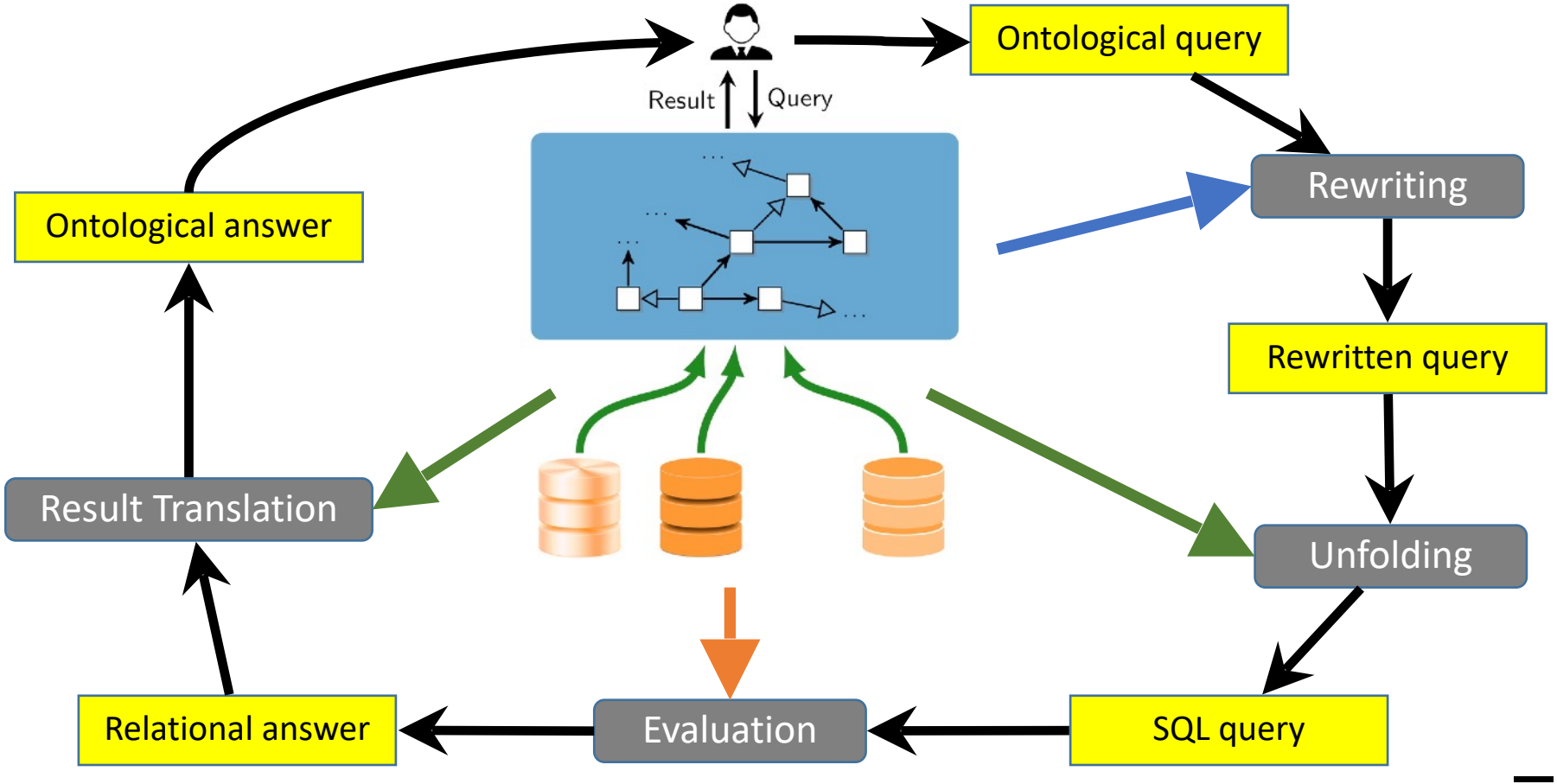
- Abstract away from data storage aspects
- Provide high-level view on data, expressed in the language of domain experts



Virtual Knowledge Graph
aka. Ontology-Based Data Access (OBDA)



VKG Approach: Query answering by rewriting



Example: data stored in a relational DB

Table Addresses

id integer	istat integer	frac_code integer	strt_code integer	label_it character varying (254)	label_de character varying (254)	num character varying	geom geometry
79520478	21008	0	8280	Via Leonardo Da Vinci 1/F	Leonardo-Da-Vinci-Strasse 1/F	1/F	01010000C...

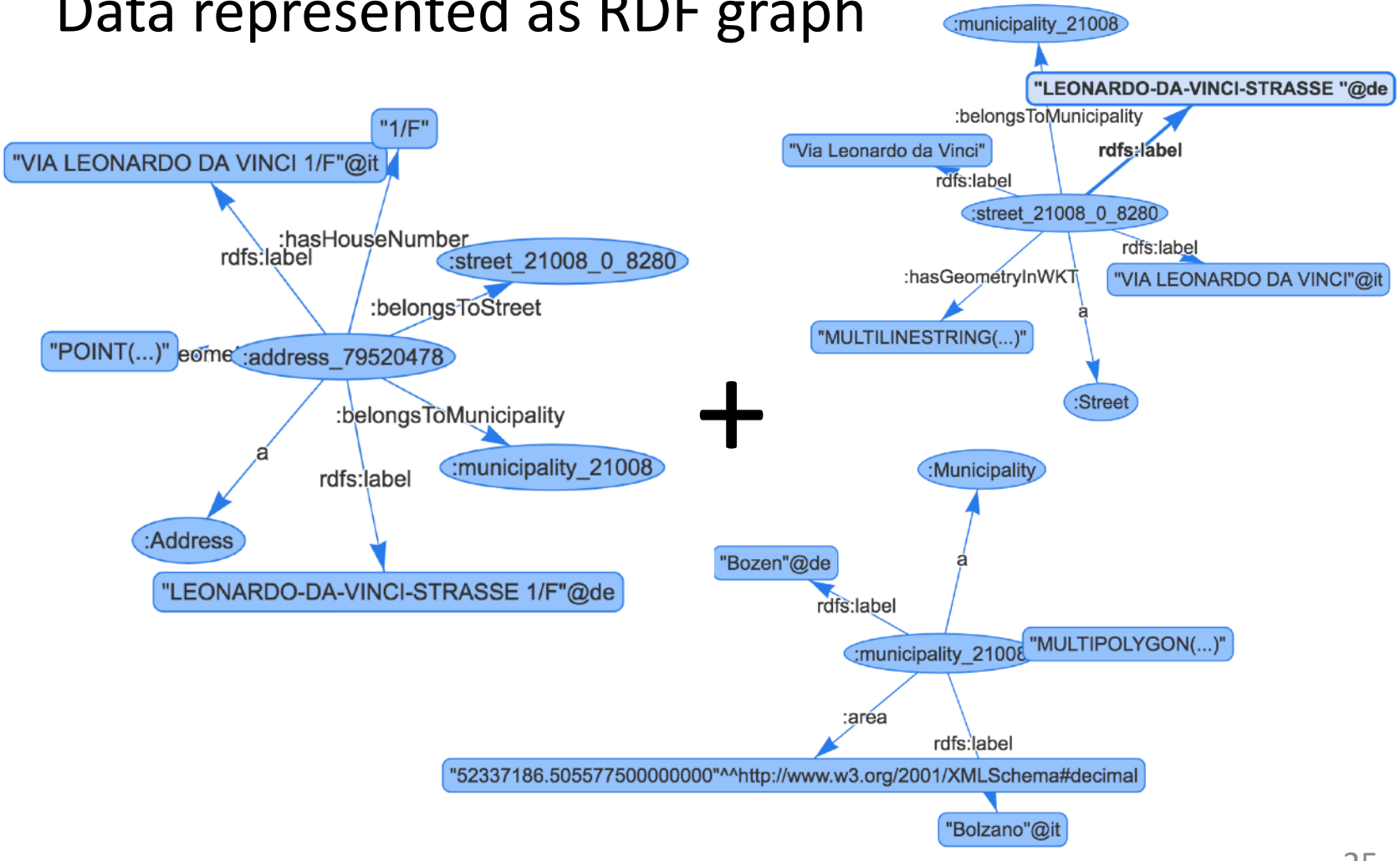
Table Municipalities

gem_id integer	istat_code integer	name_i character varying (254)	name_d character varying (254)	area numeric	geom geometry
8	21008	Bolzano	Bozen	52337186.505577500000000	...

Table Streets

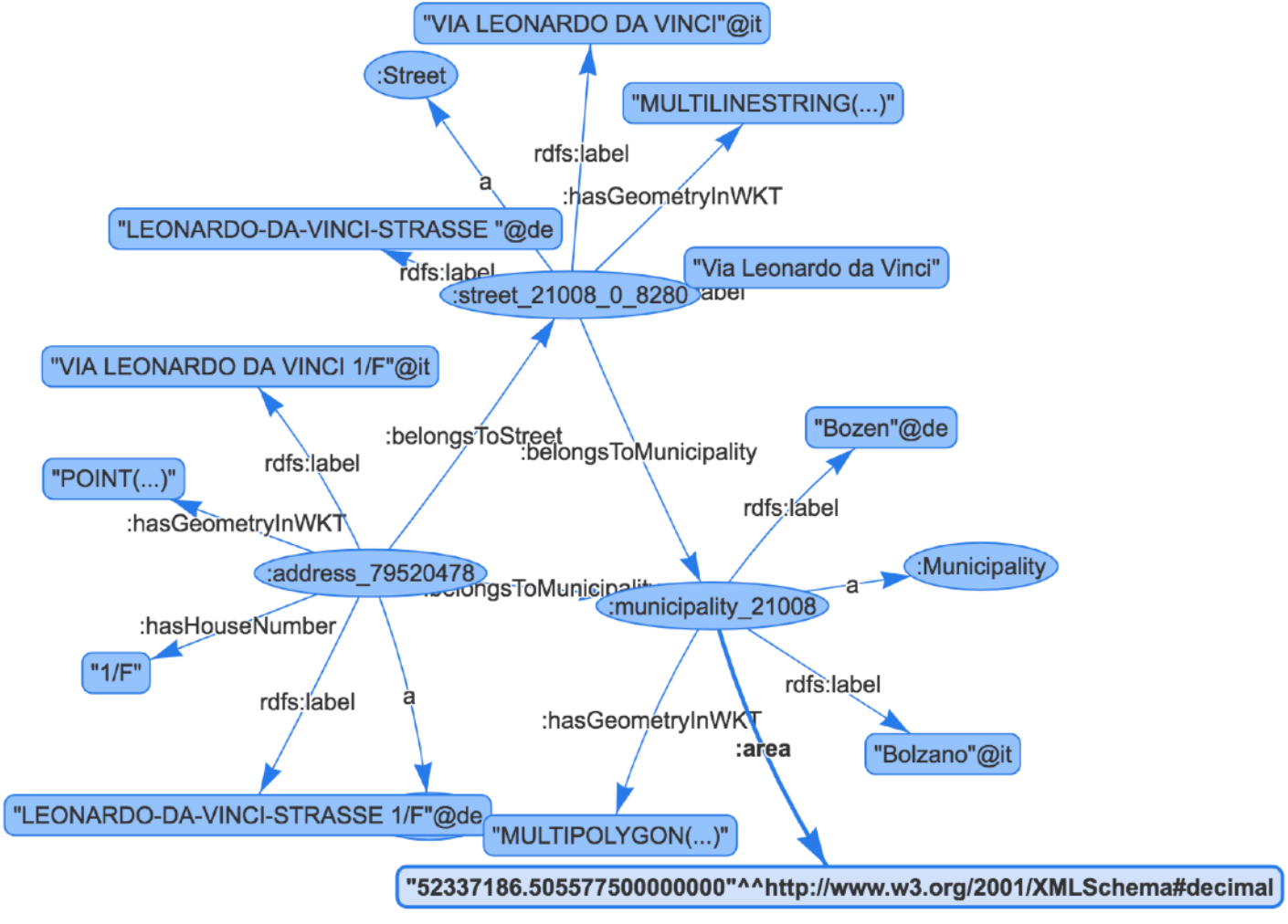
comistat integer	fraistat integer	ascot_wege integer	desc_i character varying (254)	desc_d character varying (254)	geom geometry
21008	0	8280	VIA LEONARDO DA VINCI	LEONARDO-DA-VINCI-STRASSE	01050000C...

Data represented as RDF graph



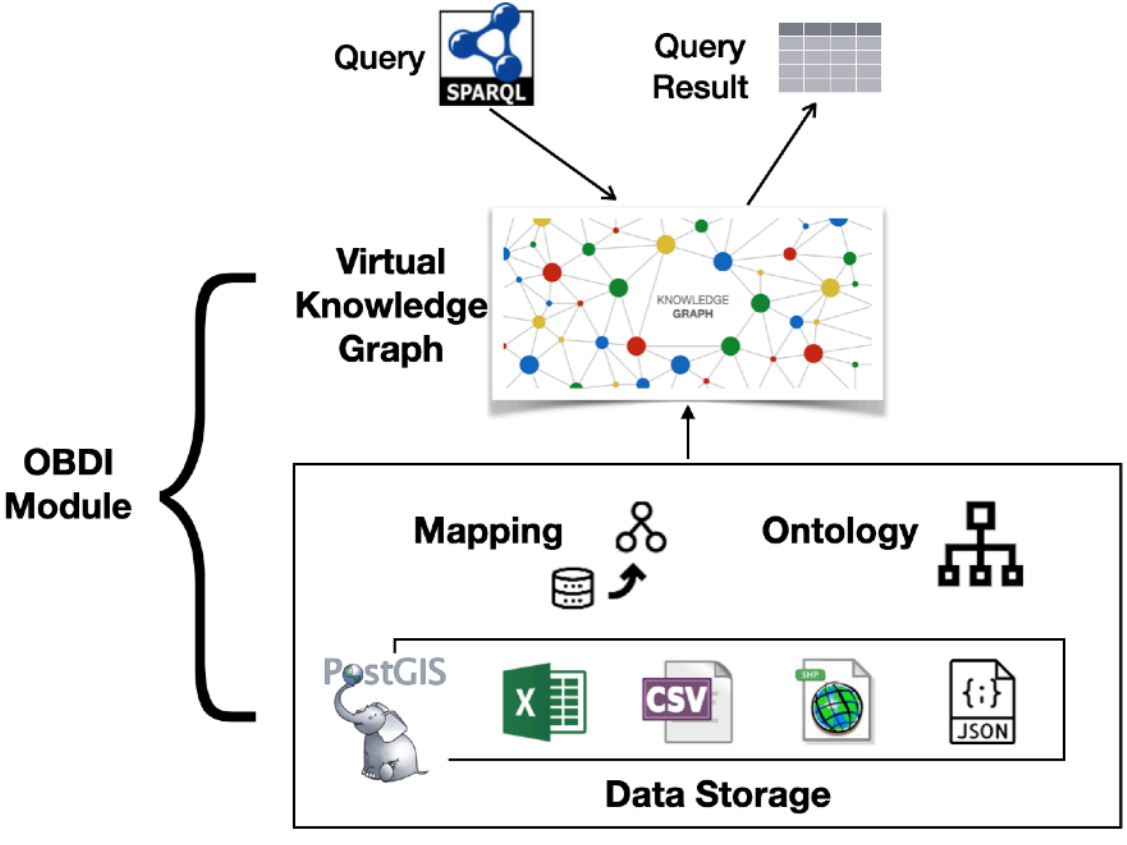
unibz

=


"VIA LEONARDO DA VINCI"@it
:Street
"MULTILINESTRING(...)"
rdfs:label
a
:hasGeometryInWKT
"LEONARDO-DA-VINCI-STRASSE "@de
rdfs:label
"Via Leonardo da Vinci"
:street_21008_0_8280
"VIA LEONARDO DA VINCI 1/F"@it
:belongsToStreet
:belongsToMunicipality
"Bozen"@de
rdfs:label
"POINT(...)"
rdfs:label
:hasGeometryInWKT
:address_79520478
:Municipality
a
:municipality_21008
:hasHouseNumber
"1/F"
rdfs:label
a
rdfs:label
:hasGeometryInWKT
:area
"Bolzano"@it
"LEONARDO-DA-VINCI-STRASSE 1/F"@de
"MULTIPOLYGON(...)"
"52337186.505577500000000"^^http://www.w3.org/2001/XMLSchema#decimal

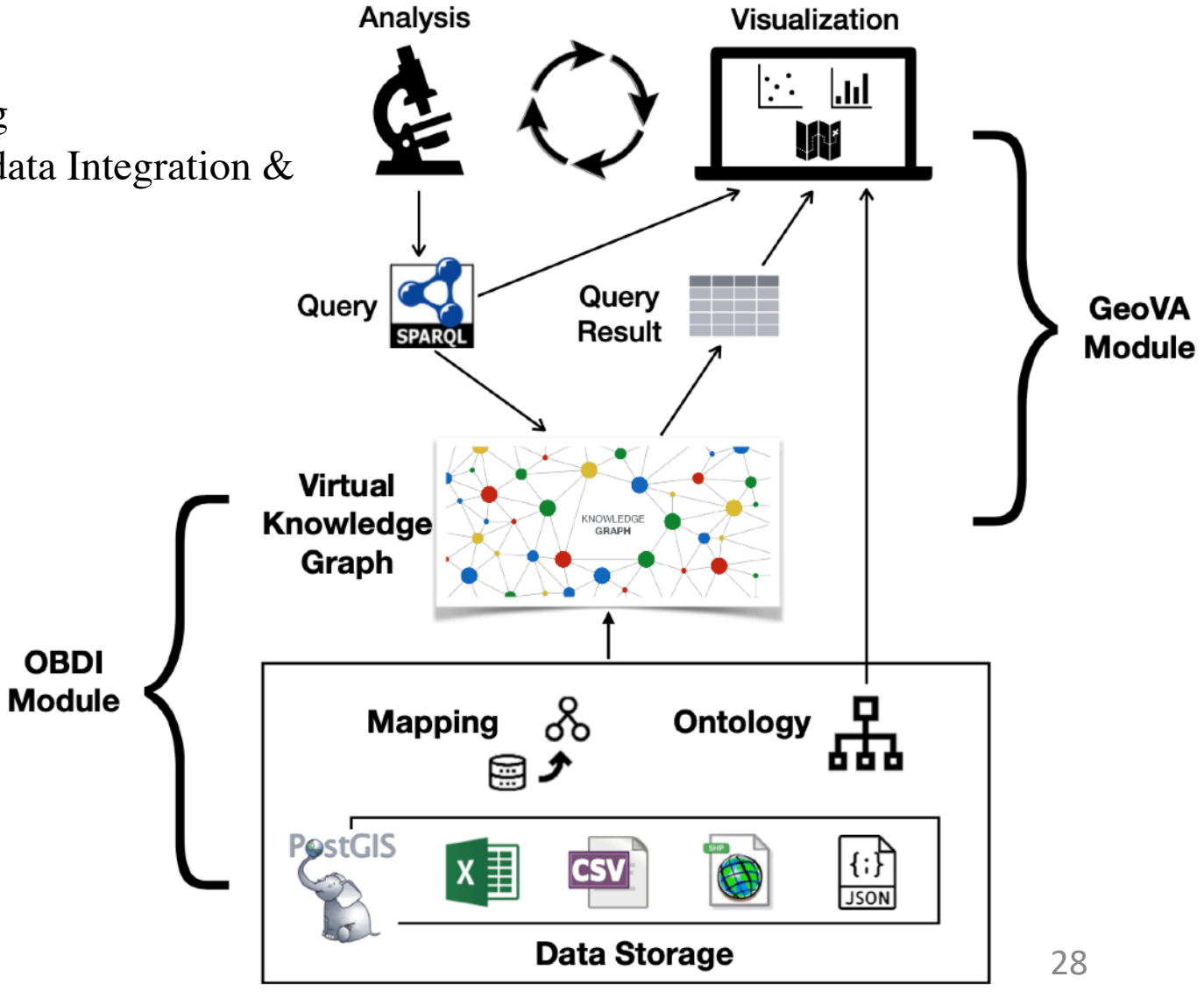
unibz

Ontology-based
Geodata Integration



unibz

A Framework Uniting
Ontology-based Geodata Integration &
Geovisual Analytics



Use cases

- Use case 1: Consistency assessment for open geodata integration
  - Understanding data quality is a prerequisite for data integration
  - Consistency is an important aspect of data quality

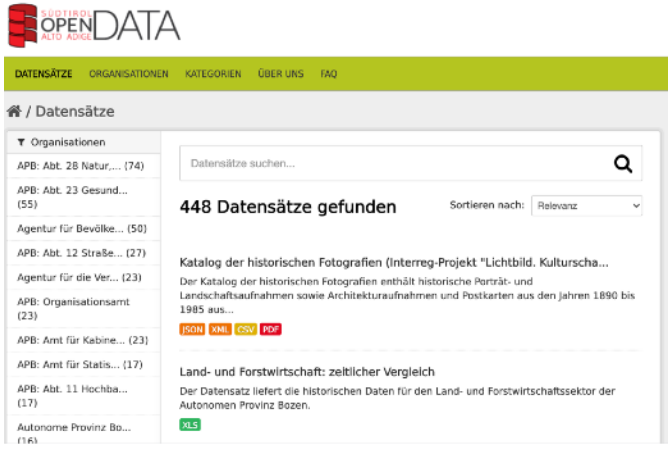
- Use case 2: Sensor data integration and analysis
  - Large amounts of sensor data help understand environment and urban dynamics

Use case 1: Consistency Assessment of Open Geodata

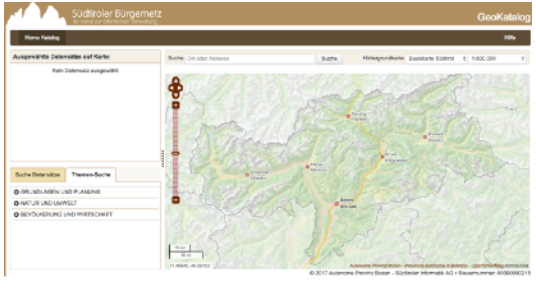
- Consistency assessment of different data sources is the crucial for producing high-quality integrated data

- *Open Government Data (OGD)* and *Volunteered Geographic Information (VGI)*

Open data portal (ODP) of South Tyrol

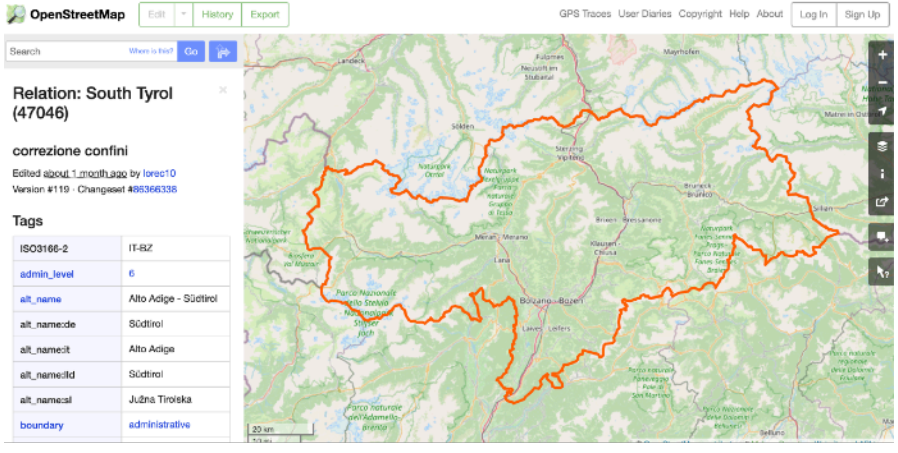


http://daten.buergernetz.bz.it/de/dataset



http://geokatalog.buergernetz.bz.it/geokatalog/#!

OpenStreetMap (OSM)



https://www.openstreetmap.org/relation/47046

Sub Region	Quick Links			
	.osm.pbf		.shp.zip	.osm.bz2
Centro	[.osm.pbf]	(244 MB)	[.shp.zip]	[.osm.bz2]
Isole	[.osm.pbf]	(138 MB)	[.shp.zip]	[.osm.bz2]
Nord-Est	[.osm.pbf]	(470 MB)	[.shp.zip]	[.osm.bz2]
Nord-Ovest	[.osm.pbf]	(398 MB)	[.shp.zip]	[.osm.bz2]
Sud	[.osm.pbf]	(231 MB)	[.shp.zip]	[.osm.bz2]

https://download.geofabrik.de/europe/italy.html

How about the inconsistency within and across these two data sources???

unibz

ODP data

dataset	description	format	number of entries
municipality	polygons of municipalities	.shp	116
street	street network	.shp	4324
address	addresses with street names and house numbers	.shp	131633
pharmacy	pharmacies with names, addresses, contact info, etc	.csv	120
organizations	organizations, (e.g., schools, museums, and offices), with names, addresses, and contact	.csv	1675 (school:1053)
healthcare	contact of sanitary offices	.csv	980
filling station	the filling stations	.shp	163

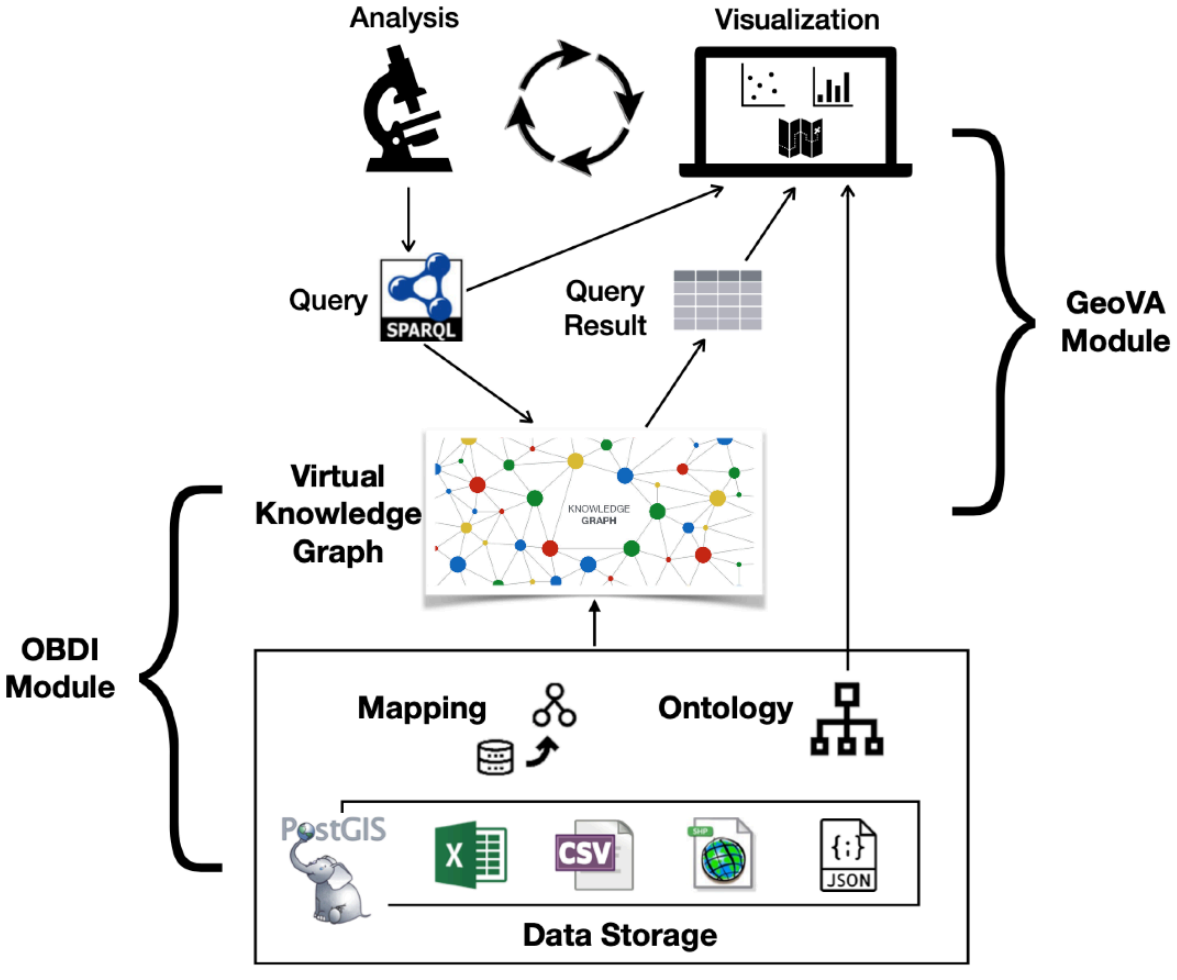
OSM data

dataset	amenity	format	number of entries
pharmacy	'pharmacy'	.shp	128 (point: 119, polygon: 9)
school	'school'/'kindergarten'	.shp	547 (point: 239, polygon: 308)
healthcare	'clinic'/'dentist'/'doctors'/ 'hospital'	.shp	90 (point:71, polygon: 19)
filling station	'fuel'	.shp	165 (point: 111, polygon: 54)

Instance-level inconsistency assessment

Schema-level inconsistency assessment

- Data structure of entities
- Schema element names
- Data types



unibz

Schema-level inconsistency: Inconsistency across ODP and OSM

- Significantly heterogeneous structures:
  - ODP: according to topics and **distributed separately in diverse formats (e.g., pdf, csv, xml, and RDF)**.
  - OSM: as **a large collection of features**, each with its geometry and a set of **flexible taggings** for different information.

- E.g., Health-related data
  - ODP: one file including all different types, like clinics, dentists, hospitals.
    additional information like address, name, telephone, doctor, and opening time.
  - OSM: points or polygons,
    can be obtained by filtering the amenity attribute with values “clinic”, … or “dentist”.

Schema-level inconsistency: Inconsistency in a single data source (ODP)

gem_id	istat_code	name_i	name_d	geom
8	21008	Bolzano	Bozen	01060000E...

(a) Table: Municipalities

comistat	fraistat	ascot_wege	desc_i	desc_d	geom
21008	0	8280	VIA LEONARDO DA VINCI	LEONARDO-DA-VINCI-STRASSE	01050000C...

(b) Table: Streets

istat	frac_code	label_it	label_de	street_code	strt_it	num	gem_it	geom
21008	0	Via Leonardo Da Vinci 1/F	Leonardo-Da-Vince Strasse 1/F	8028	VIA LEONARDO DA VINCI	1/F	BOLZANO	01010C...

(c) Table: Addresses

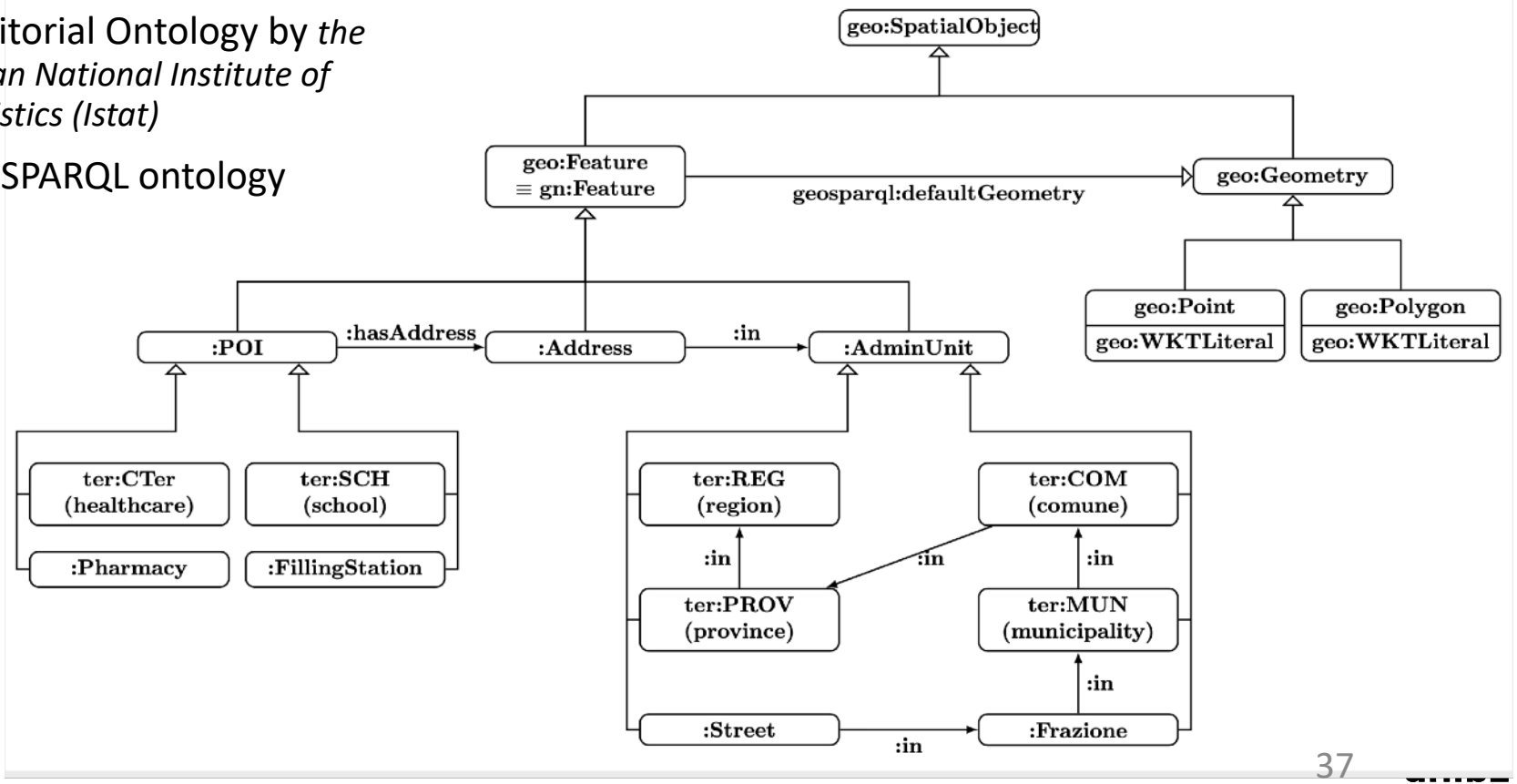
Inconsistent attribute names in ODP tables

Solutions for schema-level inconsistency

- Identified and resolved during the construction of VKG
  - by using a unified ontology and suitable mappings.

A fragment of ontology

- Territorial Ontology by *the Italian National Institute of Statistics (Istat)*
- GeoSPARQL ontology



Mapping

municipality

```
:municipality/municipality={istat_code} a :Municipality ; rdfs:label {name_i}@it , {name_d}@de ; :hasIStatCode
{istat_code}^^xsd:integer .
SELECT istat_code, name_i, name_d FROM municipalities
```

street

```
:street/municipality={comistat}/frazione={fraistat}/street={ascot_wege} a :Street ; rdfs:label {desc_d}@de , {desc_i}@it ;
:belongsToMunicipality :municipality/municipality={comistat} ; :belongsToFrazione
:frazione/municipality={comistat}/frazione={fraistat} .
SELECT * FROM roads
```

address

```
:address/{id} a :Address ; rdfs:label {label_de}@de , {label_it}@it ; :hasStreet
:street/municipality={istat}/frazione={frac_code}/street={strt_code} ; :hasPostcode {plz}^^xsd:string ; :hasHouseNumber
{num}^^xsd:string ; :hasStreetName {strt_de}@de , {strt_it}@it ; :hasStreetCode {strt_code}^^xsd:integer ; :belongsToMunicipality
:municipality/municipality={istat} ; :belongsToFrazione :frazione/municipality={istat}/frazione={frac_code} .
SELECT * FROM addresses
```

municipality_geom

```
:municipality/municipality={istat_code} :hasGeometryInWKT {wkt}^^xsd:string .
SELECT istat_code, ST_AsText(geom) AS wkt FROM municipalities
```

street_geom

```
:street/municipality={comistat}/frazione={fraistat}/street={ascot_wege} :hasGeometryInWKT {wkt}^^xsd:string .
SELECT comistat, fraistat, ascot_wege, ST_AsText(geom) AS wkt FROM roads
```

address_geom

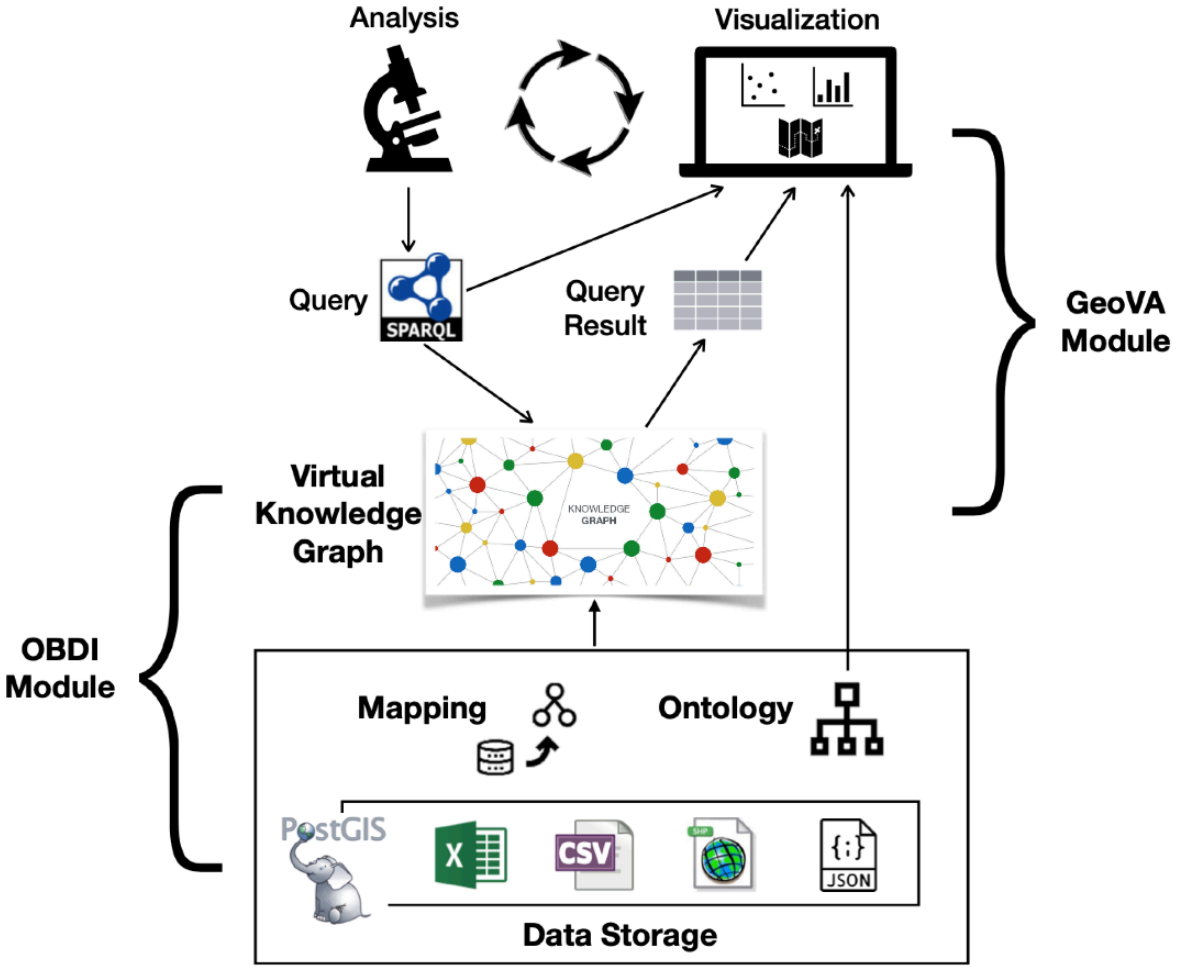
```
:address/{id} :hasGeometryInWKT {wkt}^^xsd:string .
SELECT id, ST_AsText(geom) AS wkt FROM addresses
```

Instance-level inconsistency assessment

- Thematic attributes
- Geometric attributes
- Topological properties

Schema-level inconsistency assessment

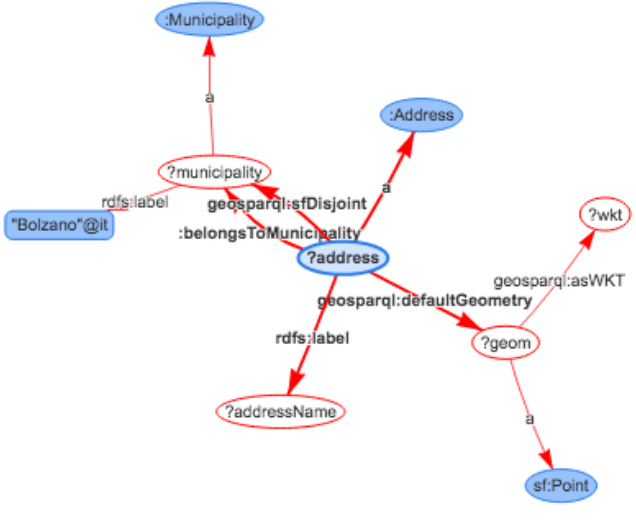
- Data structure of entities
- Schema element names
- Data types

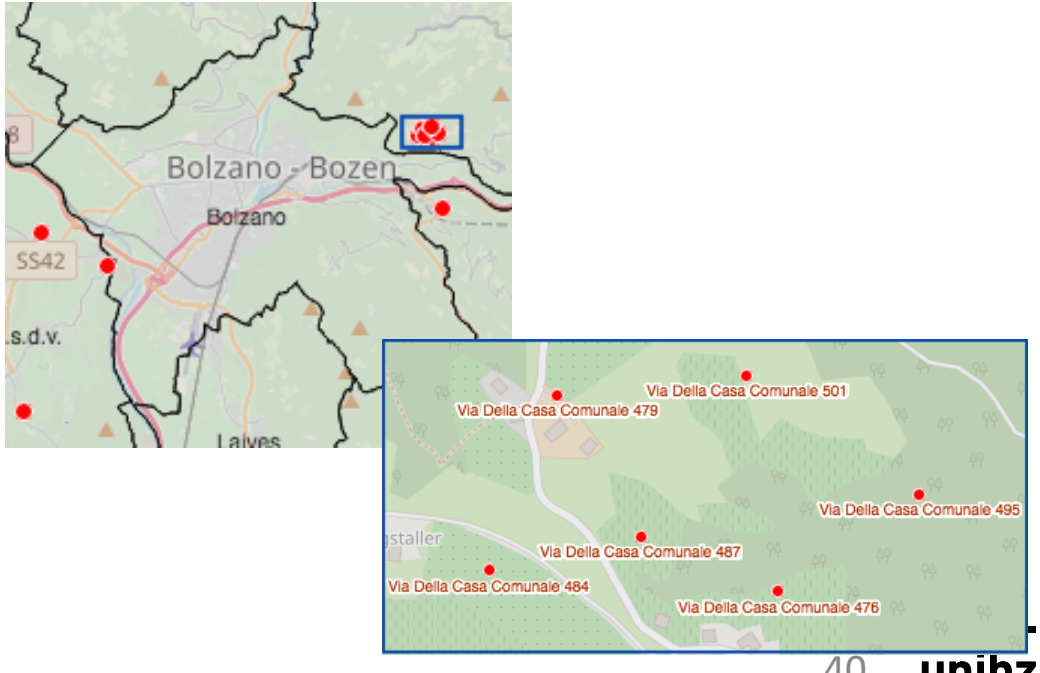


Instance-level inconsistency:
Data inconsistency in a single data source

- Inconsistent thematic and topological relations

The consistency rule: *if a feature is located within a municipality, its address should be registered in the municipality.*





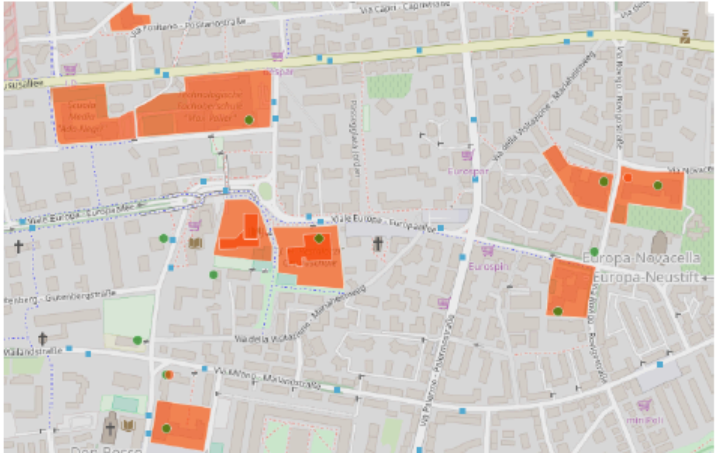
unibz

Instance-level inconsistency: Data inconsistency across ODP and OSM

- Geometric inconsistency

  The consistency rule*: if there is one school in ODP, then there must be another school in OSM (with a distance less than a predefined threshold); and vice versa*.

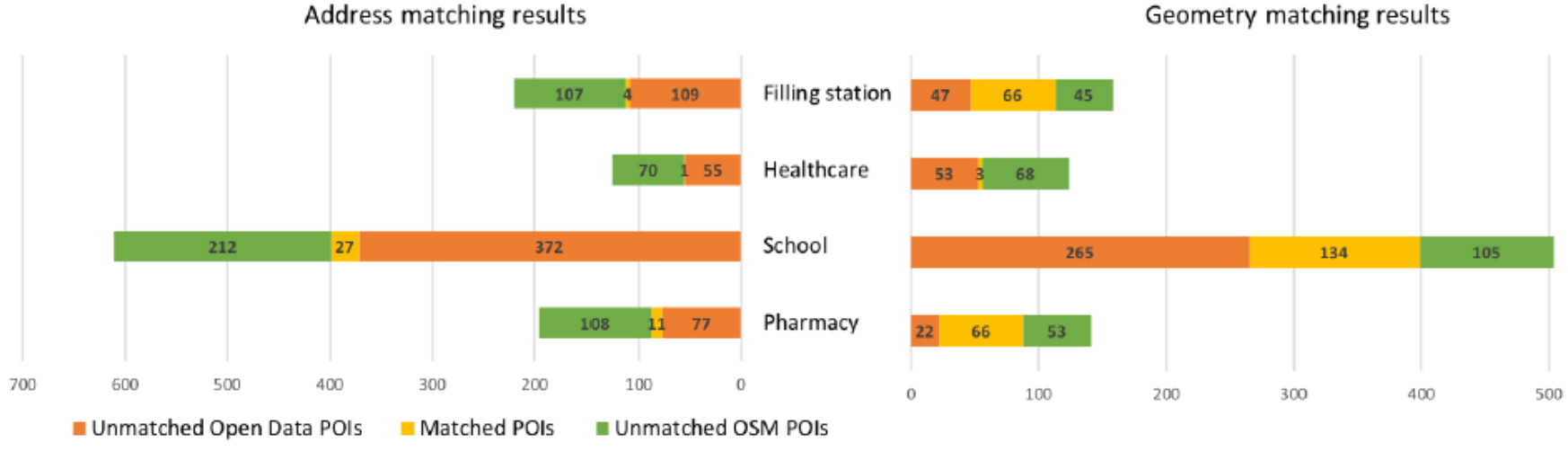
```
SELECT ?od_school ?osm_school WHERE {
  ?od_school a :School; :provenance 'OD';
  geosparql:defaultGeometry  ?od_geom.
  ?od_geom geosparql:asWKT ?od_wkt.
  ?osm_school a :School; :provenance 'OSM';
  geosparql:defaultGeometry ?osm_geom.
  ?osm_geom geosparql:asWKT ?osm_wkt.
  FILTER(ogcf:distance(?od_wkt, ?osm_wkt, 'M') < 50)
}
```



Schools of ODP and OSM

unibz

Instance-level inconsistency:
Data inconsistency across ODP and OSM

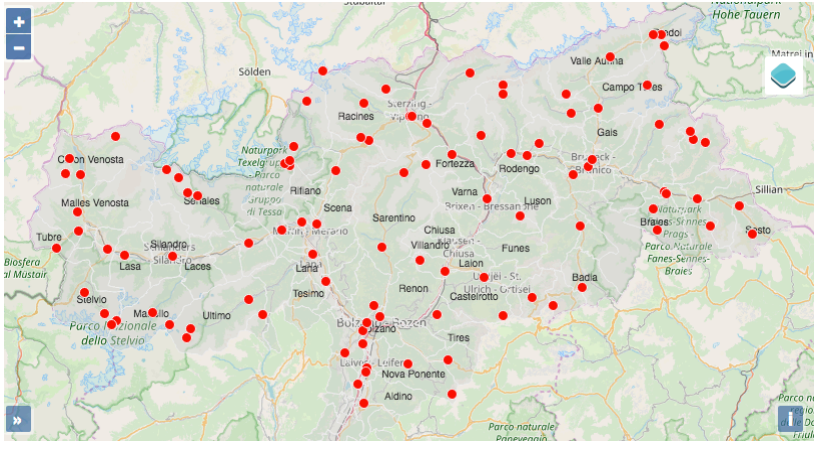


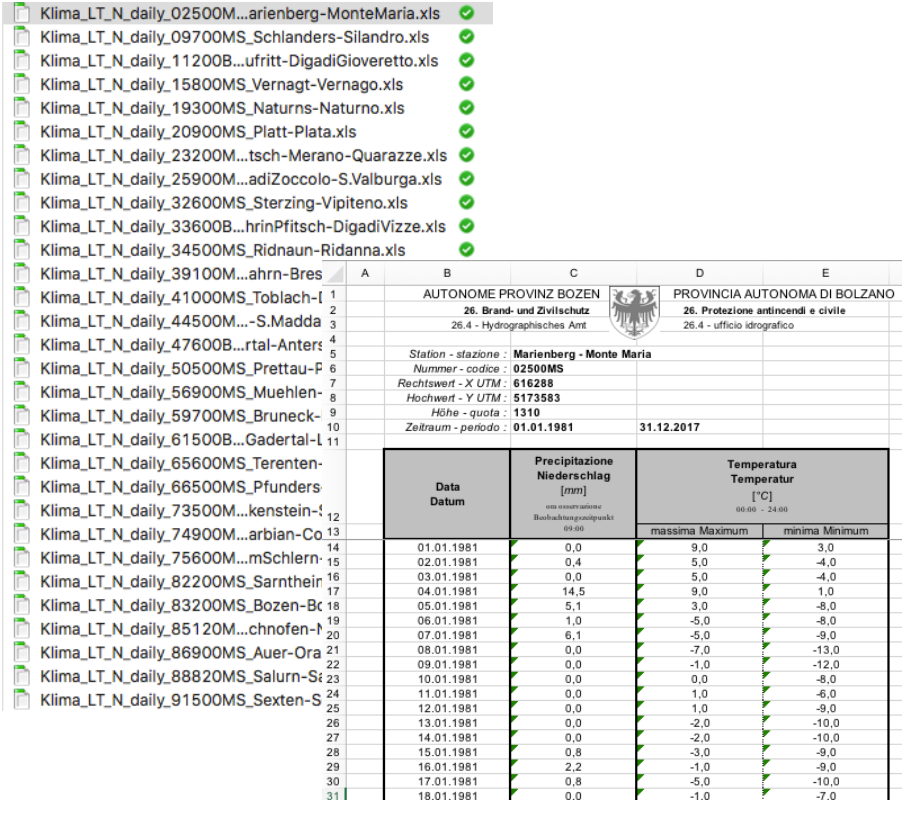
The comparison results of POIs between ODP and OSM

Use case 2: Sensor data integration and analysis

- Motivation
  - Large amounts of sensors monitoring the environment and urban dynamics
  - Understand the behaviour of complex environmental phenomena
- Meteorological and Traffic Data

Meteorological data (from the ODP of South Tyrol)





Klima_LT_N_daily_02500M...arienberg-MonteMaria.xls
Klima_LT_N_daily_09700MS_Schlanders-Silandro.xls
Klima_LT_N_daily_11200B...ufritt-DigadiGioveretto.xls
Klima_LT_N_daily_15800MS_Vernagt-Vernago.xls
Klima_LT_N_daily_19300MS_Naturns-Naturno.xls
Klima_LT_N_daily_20900MS_Platt-Plata.xls
Klima_LT_N_daily_23200M...tsch-Merano-Quarazze.xls
Klima_LT_N_daily_25900M...adiZoccolo-S.Valburga.xls
Klima_LT_N_daily_32600MS_Sterzing-Vipiteno.xls
Klima_LT_N_daily_33600B...hrinPfitsch-DigadiVizze.xls
Klima_LT_N_daily_34500MS_Ridnaun-Ridanna.xls
Klima_LT_N_daily_39100M...ahrn-Bres
Klima_LT_N_daily_41000MS_Toblach-[
Klima_LT_N_daily_44500M...-S.Madda
Klima_LT_N_daily_47600B...rtal-Anters
Klima_LT_N_daily_50500MS_Prettau-P
Klima_LT_N_daily_56900MS_Muehlen-
Klima_LT_N_daily_59700MS_Bruneck-
Klima_LT_N_daily_61500B...Gadertal-[
Klima_LT_N_daily_65600MS_Terenten-
Klima_LT_N_daily_66500MS_Pfunders
Klima_LT_N_daily_73500M...kenstein-
Klima_LT_N_daily_74900M...arbian-Co
Klima_LT_N_daily_75600M...mSchlern
Klima_LT_N_daily_82200MS_Sarntheir
Klima_LT_N_daily_83200MS_Bozen-Bc
Klima_LT_N_daily_85120M...chnofen-
Klima_LT_N_daily_86900MS_Auer-Ora
Klima_LT_N_daily_88820MS_Salurn-Sa
Klima_LT_N_daily_91500MS_Sexten-S

AUTONOME PROVINZ BOZEN
26. Brand- und Zivilschutz
26.4 - Hydrographisches Amt

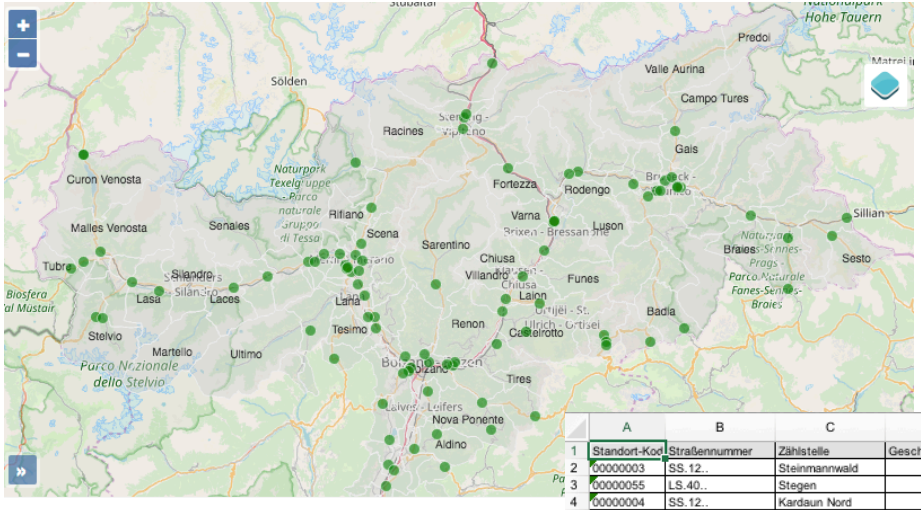
PROVINCIA AUTONOMA DI BOLZANO
26. Protezione antincendi e civile
26.4 - ufficio idrografico

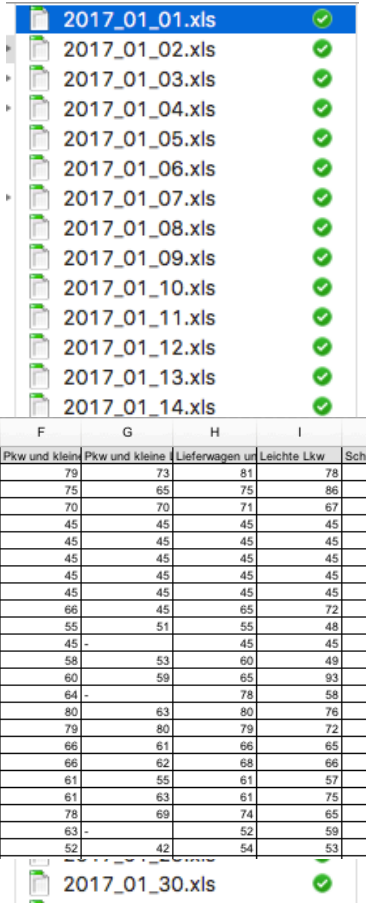
Station - stazione :	Marienberg - Monte Maria
Nummer - codice :	02500MS
Rechtswert - X UTM :	616288
Hochwert - Y UTM :	5173583
Höhe - quota :	1310
Zeitraum - periodo :	01.01.1981 – 31.12.2017

Data Datum	Precipitazione Niederschlag [mm]	Temperatura Temperatur [°C]
		massima Maximum
01.01.1981	0,0	9,0
02.01.1981	0,4	5,0
03.01.1981	0,0	5,0
04.01.1981	14,5	9,0
05.01.1981	5,1	3,0
06.01.1981	1,0	-5,0
07.01.1981	6,1	-5,0
08.01.1981	0,0	-7,0
09.01.1981	0,0	-1,0
10.01.1981	0,0	0,0
11.01.1981	0,0	1,0
12.01.1981	0,0	1,0
13.01.1981	0,0	-2,0
14.01.1981	0,0	-2,0
15.01.1981	0,8	-3,0
16.01.1981	2,2	-1,0
17.01.1981	0,8	-5,0
18.01.1981	0,0	-1,0

Data Datum	Temperatura Temperatur [°C]
	minima Minimum
01.01.1981	3,0
02.01.1981	-4,0
03.01.1981	-4,0
04.01.1981	1,0
05.01.1981	-8,0
06.01.1981	-8,0
07.01.1981	-9,0
08.01.1981	-13,0
09.01.1981	-12,0
10.01.1981	-8,0
11.01.1981	-6,0
12.01.1981	-9,0
13.01.1981	-10,0
14.01.1981	-10,0
15.01.1981	-9,0
16.01.1981	-9,0
17.01.1981	-10,0
18.01.1981	-7.0

Traffic data (from ASTAT)





	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Standort-Kod	Straßennummer	Zählstelle	Geschwindigk	Motorräder	Pkw und kleine	Pkw und kleine	Lieferwagen ur	Leichte Lkw	Schwere Lkw	Sattelzüge und	Sattelkraftfahrz	Autobusse
2	00000003	SS.12..	Steinmannwald	79	83	79	73	81	78	77	65	73	70
3	00000055	LS.40..	Stegen	75	84	75	65	75	86	70	65	76	73
4	00000004	SS.12..	Kardaun Nord	70	68	70	70	71	67	71	66	67	67
5	00000009	SS.12..	Sterzing	45	75	45	45	45	45	45	45	45	45
6	00000007	SS.12..	Vahrn	44	35	45	45	45	45	45	45	45	45
7	00000005	SS.12..	Waidbruck (Staatss	45	74	45	45	45	45	45	45	45	45
8	00000011	SS.38..	Spondinig	45	70	45	45	45	45	45	45	45	45
9	00000036	SS.238..	Marling	44	40	45	45	45	45	45	45	45	45
10	00000054	LS.24..	Waidbruck (L.S.-Ka	66	65	66	45	65	72	63	45	45	52
11	00000024	SS.44..	Thuins	55	50	55	51	55	48	43	45	48	53
12	00000053	LS.24..	Völs am Schlern	45	72	45	-	45	45	45	45	-	45
13	00000022	SS.44..	Zenoberg	58	70	58	53	60	49	52	60	51	50
14	00000018	SS.40..	Reschenpass	60	27	60	59	65	93	57	60	83	98
15	00000020	SS.42..	Sigmundskron	64	69	64	-	78	58	59	60	-	-
16	00000023	SS.44..	St. Martin in Passei	80	93	80	63	80	76	58	61	58	62
17	00000028	SS.49..	Vintl	78	69	79	80	79	72	76	79	81	64
18	00000029	SS.49..	St. Lorenzen	66	64	66	61	66	65	63	62	61	60
19	00000030	SS.49..	Bruneck Ost	66	70	66	62	68	66	65	61	66	65
20	00000034	SS.52..	Sexten	61	65	61	55	61	57	57	62	93	60
21	00000057	LS.69..	Burgstall	61	60	61	63	61	75	67	75	67	62
22	00000008	SS.12..	Mauls	78	95	78	69	74	65	65	75	75	74
23	00000042	SS.242.Dir.	Klausen	63	63	63	-	52	59	52	75	-	49
24	00000048	SS.621..	St. Georgen	52	50	52	42	54	53	49	48	46	43

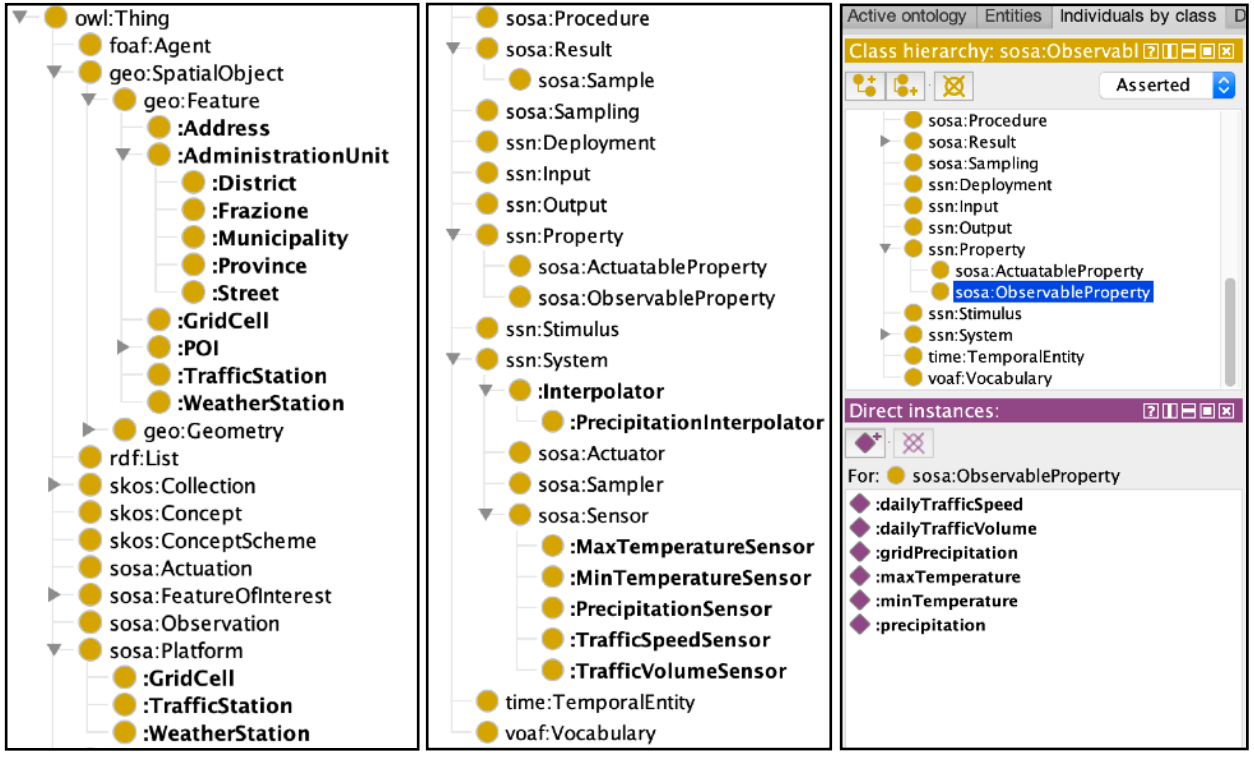
unibz

Datasets

Dataset	description	format	spatial?	temporal?	#entries	source
municipality	polygons, names (de/it), etc.	.shp	✓	-	116	ODP
meteo stations	code, name, location, etc.	.json	✓	-	84	ODP
meteo sensors	amounted station, sensor type (e.g., air temperature, precipitation)	.json	-	-	584	ODP
meteo measurements	1981 – 2017, daily min-, max-temperature, precipitation	.xls	-	✓	388,680	ODP
traffic counters	code, name, location, etc.	.shp	✓	-	75	ODP
traffic volume	daily average traffic volume in 2017	.xls	-	✓	23,381	ASTAT
traffic speed	daily average traffic speed in 2017	.xls	-	✓	23,950	ASTAT

Ontology

- Standard ontologies
  - GeoSPARQL
  - Semantic Sensor Network



unibz

Mapping

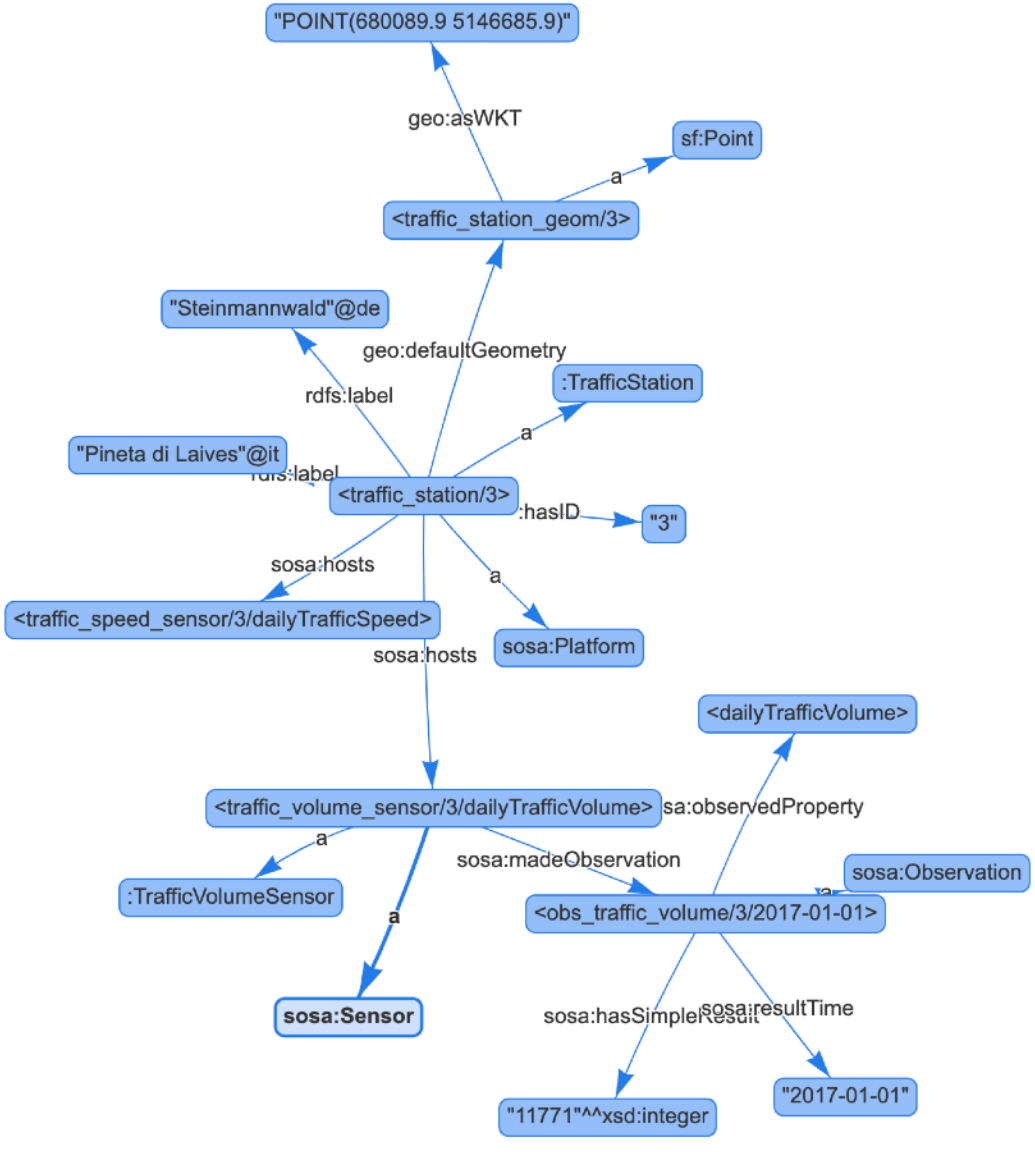
```
M_traffic_station_1:
  <traffic_station/{trst_inter}> a :TrafficStation; :hasID {trst_inter};
   rdfs:label {trst_place}@it, {trst_pla00}@de;
   :hasStreetSegmentID {trst_road_};
   :hasStreetName {trst_stree}@it,{trst_str00}@de;
   :locatesInGrid <grid/{grid_id}>;
   sosa:hosts <traffic_volume_sensor/{trst_inter}/dailyTrafficVolume>,
              <traffic_speed_sensor/{trst_inter}/dailyTrafficSpeed>.
  ← SELECT trst_inter, trst_place, trst_pla00, trst_road_,
           trst_stree, trst_str00, grid_id FROM traffic_counters
```

```
M_traffic_station_2:
  <traffic_station/{trst_inter}> geosparql:defaultGeometry
  <traffic_station_geom/{trst_inter}>.
  <traffic_station_geom/{trst_inter}> a sf:Point ; geosparql:asWKT {wkt}.
 ← SELECT trst_inter, ST_AsText(geom) AS wkt FROM traffic_counters
```

```
M_sensor_traffic_volume:
 <traffic_volume_sensor/{station_code}/dailyTrafficVolume> a
   :TrafficVolumeSensor; sosa:observes <dailyTrafficVolume>;
   sosa:madeObservation <obs_traffic_volume/{station_code}/{date}>.
 ← SELECT station_code, date FROM traffic_volume
```

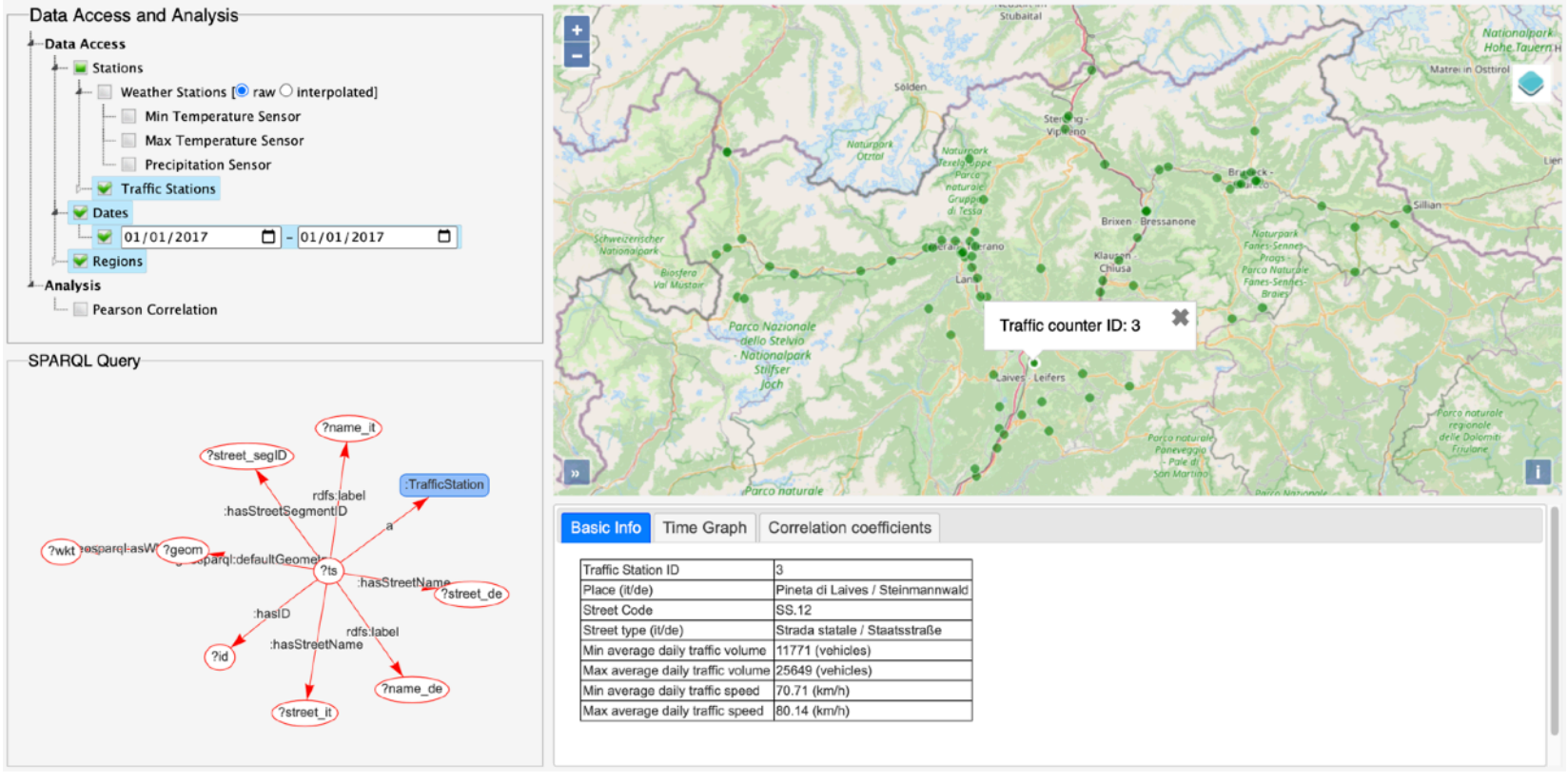
```
M_observation_traffic_volume:
  <obs_traffic_volume/{station_code}/{date}> a sosa:Observation;
   sosa:observedProperty :dailyTrafficVolume; sosa:hasSimpleResult
   {daily_volume}; sosa:resultTime {date}.
 ← SELECT station_code, date, daily_volume FROM traffic_volume
```

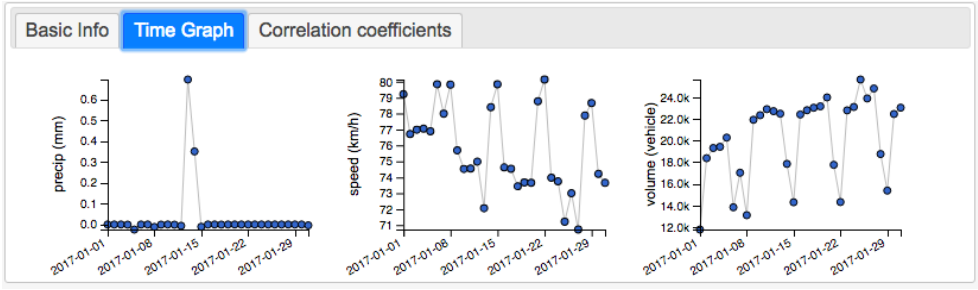
Example triples in the VKG

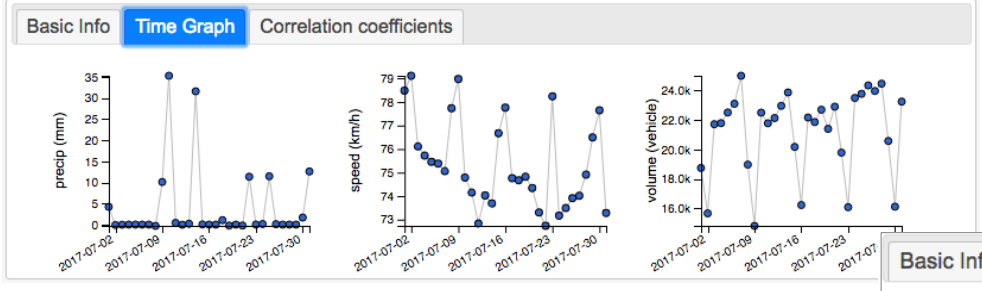


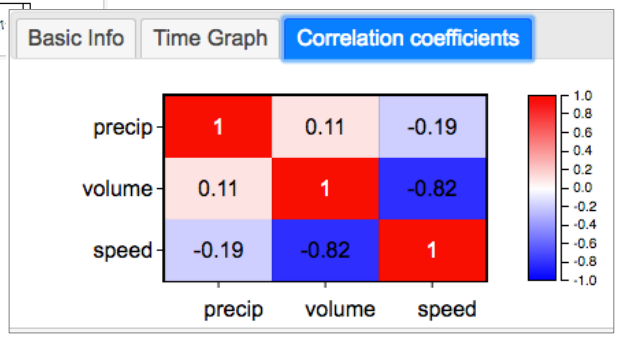
unibz

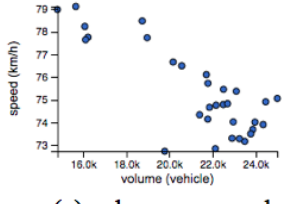
Visual Interface



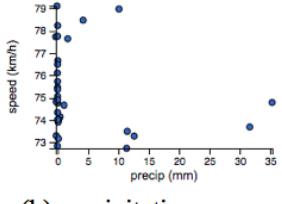




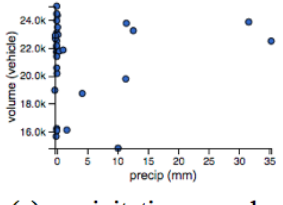




(a) volume vs. speed



(b) precipitation vs. speed



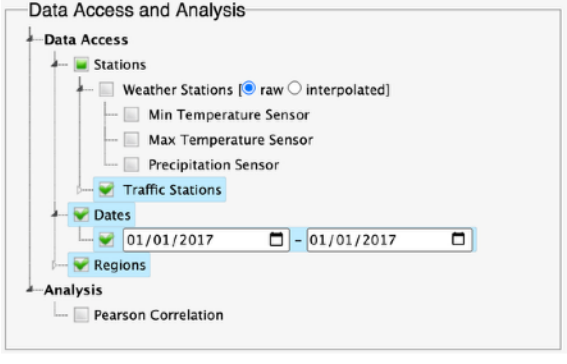
(c) precipitation vs. volume

Preliminary studies

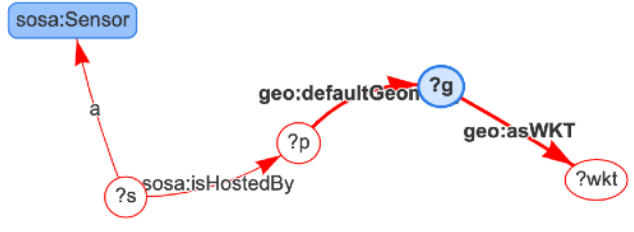
1. Exploring effectiveness:

- the formulation of sensor data analysis tasks through the visual interface

Task 1: "*Get all the sensors and their locations*."




```
SELECT * WHERE {
  ?s a sosa:Sensor .
  ?s sosa:isHostedBy ?p .
  ?p geo:defaultGeometry ?g .
  ?g geo:asWKT ?wkt .
}
```



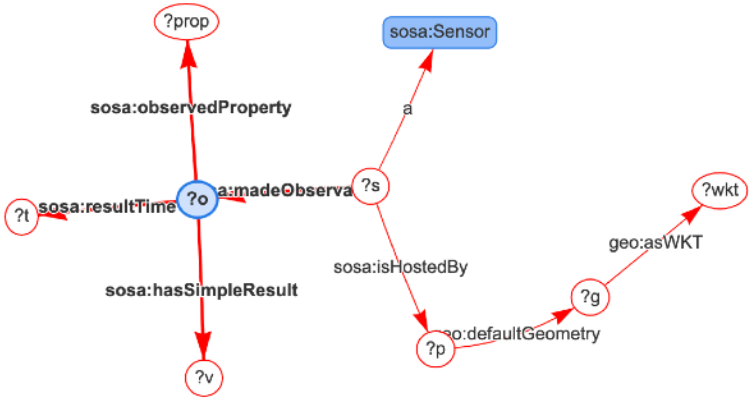
unibz

SPARQL vs. SQL

```
CONSTRUCT [s, p, g, wkt] [
p/RDF(DB_IDX_5(v0,http://ex.org/suedtirol#weather_station/{}(VARCHARToTEXT(station_code1m1)),
  http://ex.org/suedtirol#traffic_station/{}(VARCHARToTEXT(station_code3m2)),
  http://ex.org/suedtirol#weather_station/{}(VARCHARToTEXT(station_code7m3)),
  http://ex.org/suedtirol#weather_station/{}(VARCHARToTEXT(station_code11m4)),
  http://ex.org/suedtirol#traffic_station/{}(VARCHARToTEXT(station_code15m5))),IRI),
wkt/RDF(v1,http://www.opengis.net/ont/geosparql#wktLiteral),
s/RDF(DB_IDX_5(v0,http://ex.org/suedtirol#max_temperature_sensor/{}/maxTemperature(VARCHARToTEXT(station_code1m1
  http://ex.org/suedtirol#traffic_volume_sensor/{}/dailyTrafficVolume(VARCHARToTEXT(station_code3m2)),
  http://ex.org/suedtirol#min_temperature_sensor/{}/minTemperature(VARCHARToTEXT(station_code7m3)),
  http://ex.org/suedtirol#precipitation_sensor/{}/precipitation(VARCHARToTEXT(station_code11m4)),
  http://ex.org/suedtirol#traffic_speed_sensor/{}/dailyTrafficSpeed(VARCHARToTEXT(station_code15m5))),IRI),
g/RDF(DB_IDX_5(v0,http://ex.org/suedtirol#weather_station_geom/{}(VARCHARToTEXT(station_code1m1)),
  http://ex.org/suedtirol#traffic_station_geom/{}(VARCHARToTEXT(station_code3m2)),
  http://ex.org/suedtirol#weather_station_geom/{}(VARCHARToTEXT(station_code7m3)),
  http://ex.org/suedtirol#weather_station_geom/{}(VARCHARToTEXT(station_code11m4)),
  http://ex.org/suedtirol#traffic_station_geom/{}(VARCHARToTEXT(station_code15m5))),IRI)]


NATIVE [station_code11m4, station_code15m5, station_code1m1, station_code3m2, station_code7m3, v0, v1]

SELECT v26."station_code11m4" AS "station_code11m4", v26."station_code15m5" AS "station_code15m5",
 v26."station_code1m1" AS "station_code1m1", v26."station_code3m2" AS "station_code3m2",
 v26."station_code7m3" AS "station_code7m3", v26."v0" AS "v0", v26."v1" AS "v1"
FROM (SELECT CAST(NULL AS VARCHAR) AS "station_code11m4", CAST(NULL AS VARCHAR) AS "station_code15m5",
  v4."station_code1m1" AS "station_code1m1", CAST(NULL AS VARCHAR) AS "station_code3m2",
  CAST(NULL AS VARCHAR) AS "station_code7m3", 0 AS "v0", CAST(ST_ASTEXT(v4."geom4m20") AS TEXT) AS "v1"
 FROM (SELECT DISTINCT v2."geom" AS "geom4m20", v1."station_code" AS "station_code1m1"
  FROM "meteo_measurements" v1, "meteo_stations" v2
  WHERE (v2."geom" IS NOT NULL AND v1."station_code" = v2."scode")
  ) v4
  UNION ALL
  SELECT CAST(NULL AS VARCHAR) AS "station_code11m4", CAST(NULL AS VARCHAR) AS "station_code15m5",
   CAST(NULL AS VARCHAR) AS "station_code1m1", v9."station_code3m2" AS "station_code3m2",
   CAST(NULL AS VARCHAR) AS "station_code7m3", 1 AS "v0", CAST(ST_ASTEXT(v9."geom4m20") AS TEXT) AS "v1"
  FROM (SELECT DISTINCT v7."geom" AS "geom4m20", v6."station_code" AS "station_code3m2"
   FROM "traffic_volume" v6, "traffic_counters" v7
   WHERE (v7."geom" IS NOT NULL AND v6."station_code" = v7."trst_inter")
   ) v9
  UNION ALL
  SELECT CAST(NULL AS VARCHAR) AS "station_code11m4", CAST(NULL AS VARCHAR) AS "station_code15m5",
   CAST(NULL AS VARCHAR) AS "station_code1m1", CAST(NULL AS VARCHAR) AS "station_code3m2",
   v14."station_code7m3" AS "station_code7m3", 2 AS "v0", CAST(ST_ASTEXT(v14."geom4m20") AS TEXT) AS "v1"
  FROM (SELECT DISTINCT v12."geom" AS "geom4m20", v11."station_code" AS "station_code7m3"
   FROM "meteo_measurements" v11, "meteo_stations" v12
   WHERE (v12."geom" IS NOT NULL AND v11."station_code" = v12."scode")
   ) v14
  UNION ALL
  SELECT v19."station_code11m4" AS "station_code11m4", CAST(NULL AS VARCHAR) AS "station_code15m5",
   CAST(NULL AS VARCHAR) AS "station_code1m1", CAST(NULL AS VARCHAR) AS "station_code3m2",
   CAST(NULL AS VARCHAR) AS "station_code7m3", 3 AS "v0", CAST(ST_ASTEXT(v19."geom4m20") AS TEXT) AS "v1"
  FROM (SELECT DISTINCT v17."geom" AS "geom4m20", v16."station_code" AS "station_code11m4"
   FROM "meteo_measurements" v16, "meteo_stations" v17
   WHERE (v17."geom" IS NOT NULL AND v16."station_code" = v17."scode")
   ) v19
  UNION ALL
  SELECT CAST(NULL AS VARCHAR) AS "station_code11m4", v24."station_code15m5" AS "station_code15m5",
   CAST(NULL AS VARCHAR) AS "station_code1m1", CAST(NULL AS VARCHAR) AS "station_code3m2",
   CAST(NULL AS VARCHAR) AS "station_code7m3", 4 AS "v0", CAST(ST_ASTEXT(v24."geom4m20") AS TEXT) AS "v1"
  FROM (SELECT DISTINCT v22."geom" AS "geom4m20", v21."station_code" AS "station_code15m5"
   FROM "traffic_speed" v21, "traffic_counters" v22
   WHERE (v22."geom" IS NOT NULL AND v21."station_code" = v22."trst_inter")
   ) v24
  ) v26
```

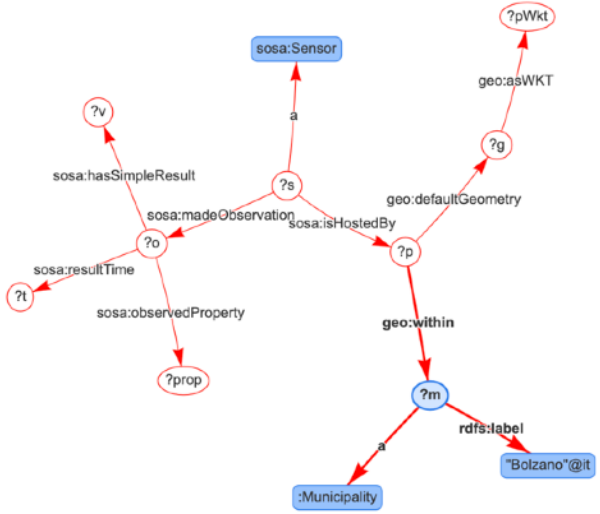
Task 2: *"Get all the sensors, their locations, and observations on 1 January 2017."*

```
SELECT * WHERE {
  ?s a sosa:Sensor .
  ?s sosa:isHostedBy ?p.
  ?p geo:defaultGeometry ?g .
  ?g geo:asWKT ?wkt .

  ?s sosa:madeObservation ?o .
  ?o sosa:hasSimpleResult ?v .
  ?o sosa:resultTime ?t .
  FILTER (?t = '2017-01-01'^^xsd:date)
```



Task 3: "*Get all the sensors, their locations, and observations in the municipality of Bolzano on 1 January 2017.*"

```
SELECT * WHERE {
  ?s a sosa:Sensor .
  ?s sosa:isHostedBy ?p.
  ?p rdfs:label ?pName .
  ?p geo:defaultGeometry ?g .
  ?g geo:asWKT ?pWkt .

  ?s sosa:madeObservation ?o .
  ?o sosa:observedProperty ?prop .
  ?o sosa:hasSimpleResult ?v .
  ?o sosa:resultTime ?t .
  FILTER (?t = '2017-01-01'^^xsd:date)

  ?m a :Municipality .
  ?m rdfs:label 'Bolzano'@it.
  ?p geo:within ?m.
}
```



Impacts

- First presentation at the 9th Workshop of
  “Computer Science Research Meets Business” on GIS and Location-based Services,
  23 November 2017, Unibz.

  Attendants showed strong interests:

  (1) Südtiroler Informatik AG (SIAG), who is managing the OpenDataPortal,

  (2) ASTAT, who is in charge of the local traffic data,

  (3) NOI Techpark, a local service provider for companies

  (4) R3 GIS, an SME specialized in the development of GIS technology.

- Several follow-up meetings, including dedicated demos and a hackthon

Following up Projects

1. *IDEE*: Data Integration for Energy Efficiency is a 3-year project supported by European Regional Development Fund (ERDF). The consortium consists of
    - unibz: geodata integration solution provider
    - Alperia: energy consumption data provider
    - R3 GIS: GIS infrastructure provider
    - The city of Merano: the main use-case partner (requirements and data)
2. *Open Data Hub-Virtual Knowledge Graph*
    - Joint project between NOI techpark and Ontopic
    - Extend the South Tyrolean OpenDataHub with a Knowledge Graph Interface.

Conclusions

- A framework uniting OBDI an GeoVA to integrate and visually analyse complex geospatial data is a very promising approach.
- Demonstrated by two use cases:
  - consistency of open geospatial data.
  - integration and visual analysis the sensor data.
- Future Work:
  - (semi)-automate the VKG construction
  - more analytics functions

Thank you for your attention!

Reference:

- Ding, L., Xiao, G., Calvanese, D., Meng, L. Consistency assessment for open geodata integration: an ontology-based approach. *Geoinformatica* (2019). https://doi.org/10.1007/s10707-019-00384-9
- Ding, L., Xiao, G., Calvanese, D., Meng, L. A Framework Uniting Ontology-based Geodata Integration and Geovisual Analytics. *Submitted to ISPRS IJGI.*