



Creative Commons Attribution 4.0 International (CC BY 4.0)



Co-financed by the Connecting Europe Facility of the European Union



From data to knowledge: the role of semantic technologies

Andrea Giovanni Nuzzolese

Consiglio Nazionale delle Ricerche
Istituto di Scienze e Tecnologie della Cognizione
Semantic Technology Laboratory

23/03/2023

Open data value

Open data is a tremendous resource for transparency and democratic control, participation, self-empowerment and economic growth.

It helps citizens and organisations to make better decisions, or enable you to be more active in society.



Any real value for citizens?

Hazardous substances and other chemical determinants in lakes - aggregated data

NationalStationID	Year	DeterminandHazSubs	CASNumber	NumberOfSamples	NumberOfSamplesBelLOQ	LimitOfQuantification	Minimum	Mean	Maximum	Median	StandardDeviation	SampleDepth	NoOfSubsites	MethodSubsitesSpatialAggreg	Remarks
22-SGLA0011	2011	105	7440-38-2	5	3	0	5	0	25	0	49	0	9	0	25
22-SGLA0011	2011	121	7440-43-9	5	5	0	5	0	25	0	25	0	25	0	25
22-SGLA0011	2011	145	7440-47-3	5	5	2	1	1	1	1	0	0			
22-SGLA0011	2011	132	7439-89-6	5	0	5	5	11	15	12	3	8079	0		
22-SGLA0011	2011	155	7439-96-5	5	5	1	0	5	0	5	0	5	0	5	0
22-SGLA0011	2011	161	7439-97-6	5	5	0	1	0	5	0	5	0	5	0	5
22-SGLA0011	2011	173	7440-02-0	5	5	2	1	1	1	1	0	0			
22-SGLA0011	2011	172	7439-92-1	5	5	1	0	5	0	5	0	5	0	5	0
22-SGLA0011	2011	131	7440-50-8	5	5	5	2	5	2	5	2	5	2	5	0
22-SGLA0011	2011	144	7440-66-6	5	4	10	5	8	20	5	6	7082	0		
22-SGLA0011	2011	105	7440-38-2	5	3	0	5	0	25	0	49	0	9	0	25

Open data retrieved from dati.gov.it

Any real value for citizens?

Hazardous substances and other chemical determinants in lakes - aggregated data

NationalStationID	Year	DeterminandHazSubs	CASNumber	NumberOfSamples	NumberOfSamplesBelLOQ	LimitOfQuantification	Minimum	Mean	Maximum	Median	StandardDeviation	SampleDepth	NoOfSubsites	MethodSubsitesSpatialAggreg	Remarks
22-SGLA0011	2011	105	7440-38-2	5	3	0	5	0	25	0	49	0	9	0	25
22-SGLA0011	2011	121	7440-43-9	5	5	0	5	0	25	0	25	0	25	0	25
22-SGLA0011	2011	145	7440-47-3	5	5	2	1	1	1	1	0	0			
22-SGLA0011	2011	132	7439-89-6	5	0	5	5	11	15	12	3	8079	0		
22-SGLA0011	2011	155	7439-96-5	5	5	1	0	5	0	5	0	5	0	5	0
22-SGLA0011	2011	161	7439-97-6	5	5	0	1	0	5	0	5	0	5	0	5
22-SGLA0011	2011	173	7440-02-0	5	5	2	1	1	1	1	0	0			
22-SGLA0011	2011	172	7439-92-1	5	5	1	0	5	0	5	0	5	0	5	0
22-SGLA0011	2011	131	7440-50-8	5	5	5	2	5	2	5	2	5	2	5	0
22-SGLA0011	2011	144	7440-66-6	5	4	10	5	8	20	5	6	7082	0		
22-SGLA0011	2011	105	7440-38-2	5	3	0	5	0	25	0	49	0	9	0	25

Open data retrieved from dati.gov.it



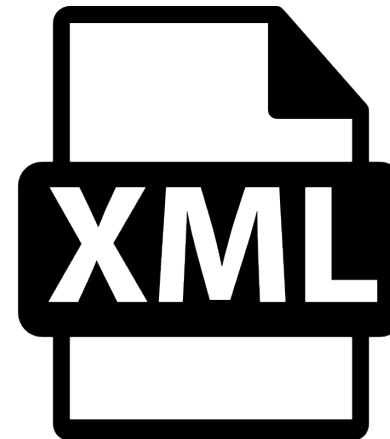


No open data without open formats

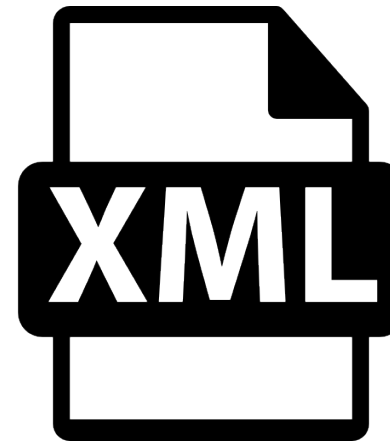
“The format of digital data is defined as “open” when its syntax, semantics, operational context and mode of use are made public through comprehensive documentation.”

Source: <https://www.agid.gov.it/it/dati/formati-aperti>

(Some) Open data formats



(Some) Open data formats



Open data as

Knowledge soup: Knowledge can be extracted from a plethora of different data sources, which are inherently heterogeneous with respect to syntax and semantics



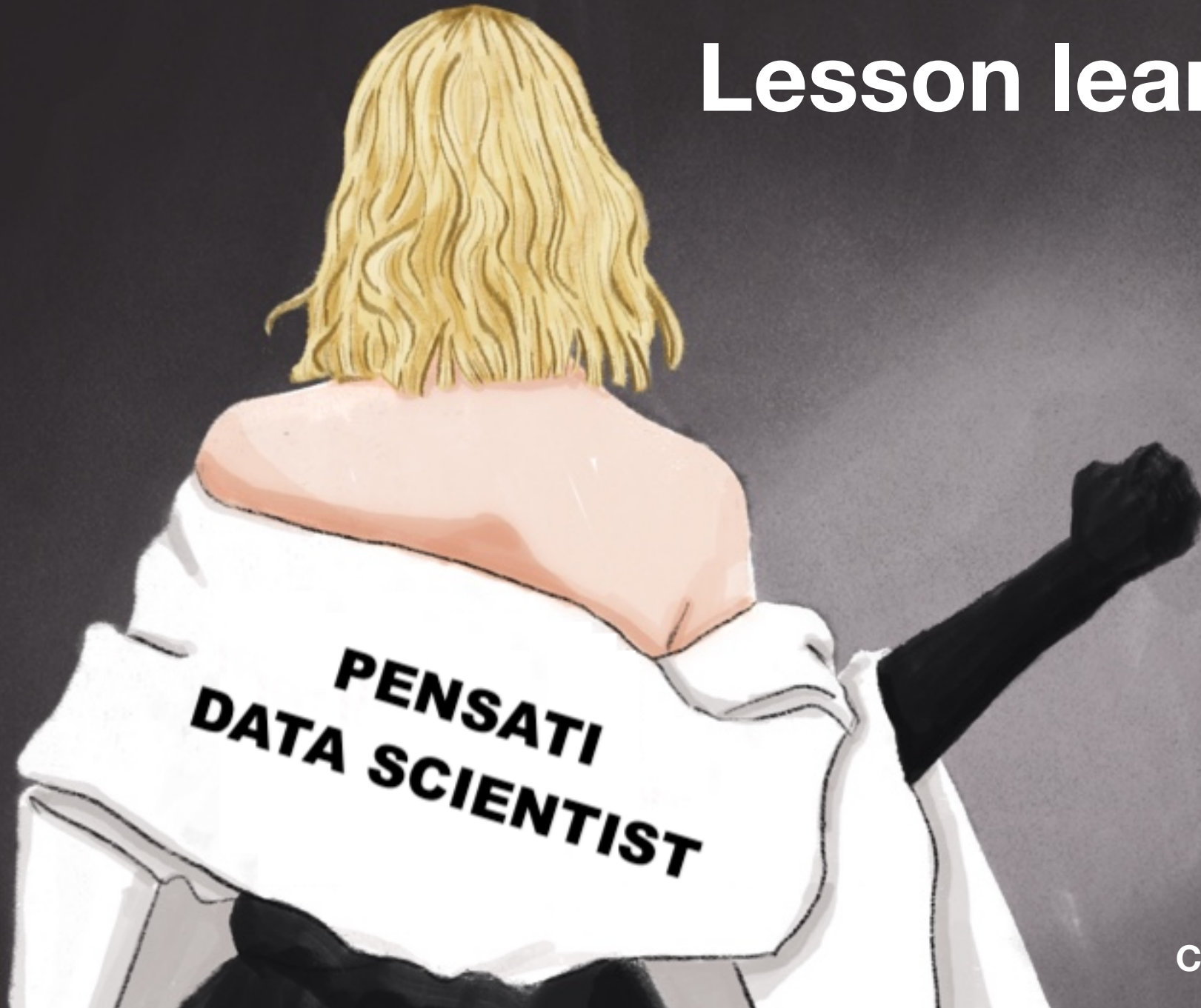
Open data as

Big data: (Heterogeneous) Data sources contain billions of data that should be processed. Hence, how do we get only contextual data which are relevant to my goals?



The Emu in the Sky, a well-known Aboriginal Astronomical constellation

Lesson learnt so far



Credits: Sonia Giampaolo

Open data FAIRness



Findable



data.europa.eu



dati.gov.it



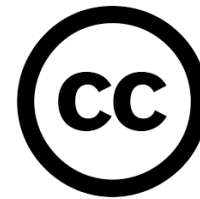
Accessible



Interoperable



Reusable



Open data FAIRness



Findable



Accessible



Interoperable



Reusable



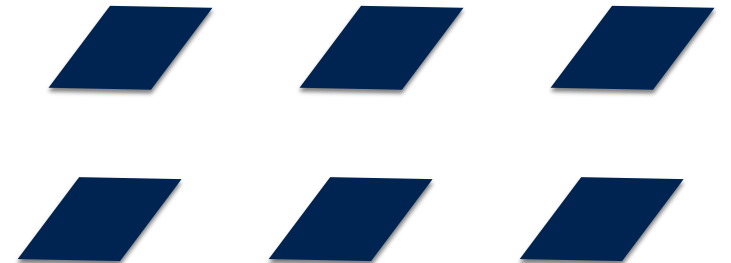
**Lesson learnt
(contd.)**

A little semantics goes a long way



Jim Hendler's hypothesis <https://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html>

The open data is a flat, boring world devoid
of meaning



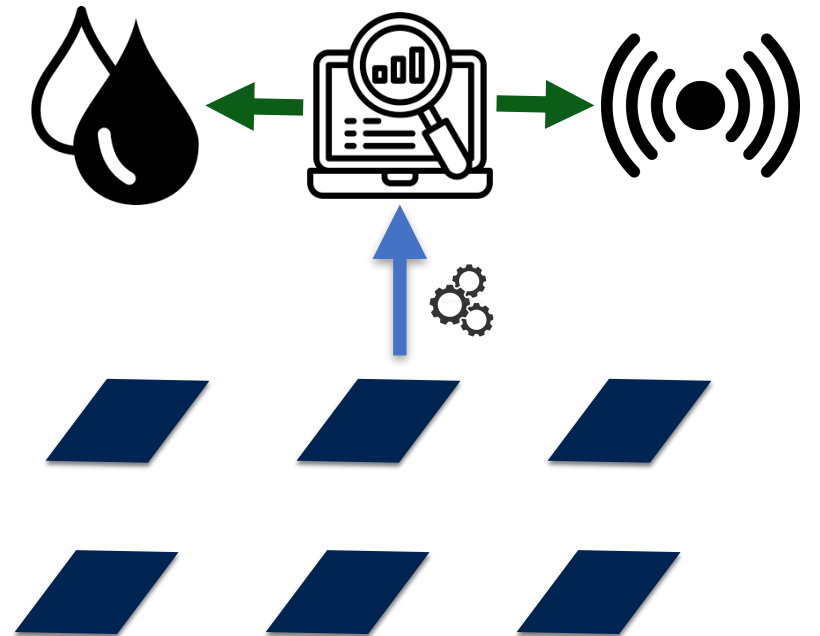
A little semantics goes a long way



Jim Hendler's hypothesis <https://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html>

This is a pity, as data describe real objects with meanings and relationships between them

The open data is a flat, boring world devoid of meaning



A little semantics goes a long way



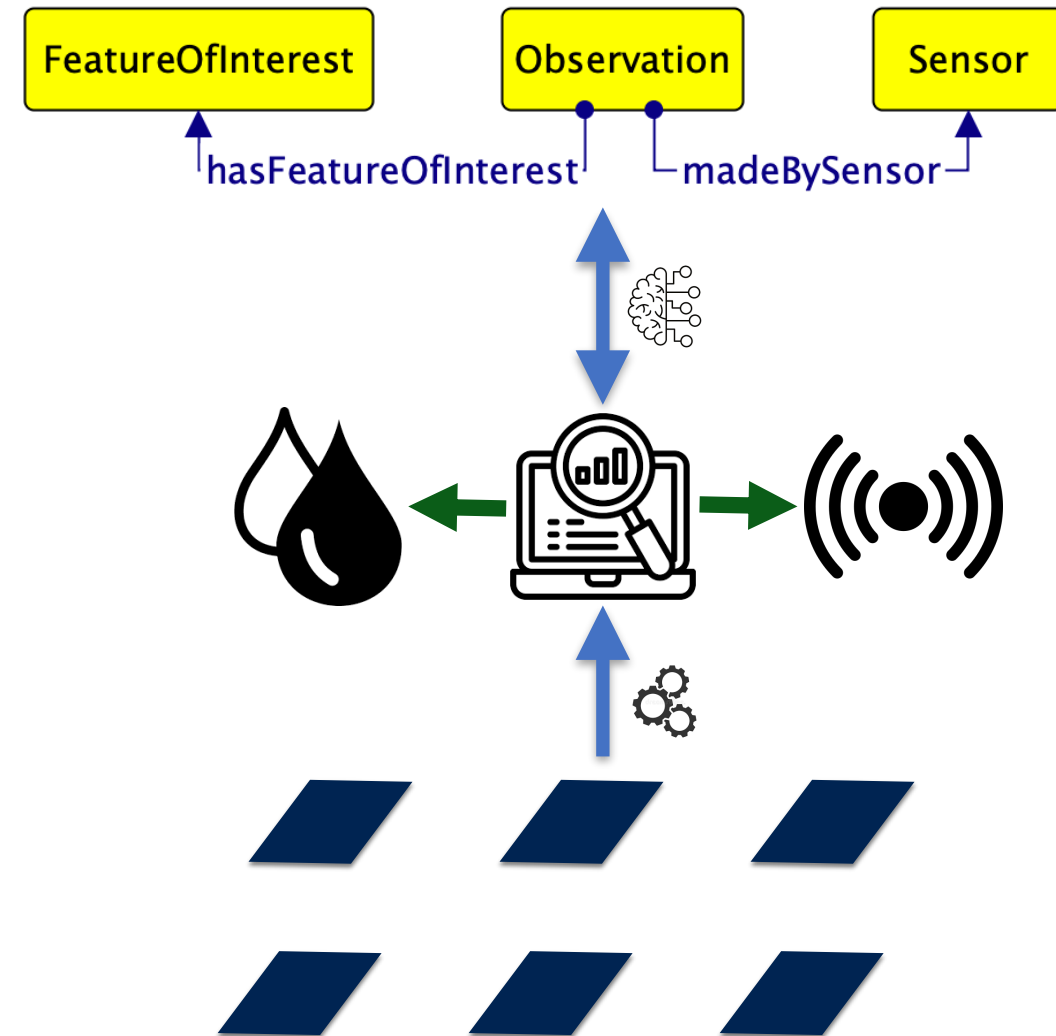
Jim Hendler's hypothesis <https://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html>

From open data to linked open data:

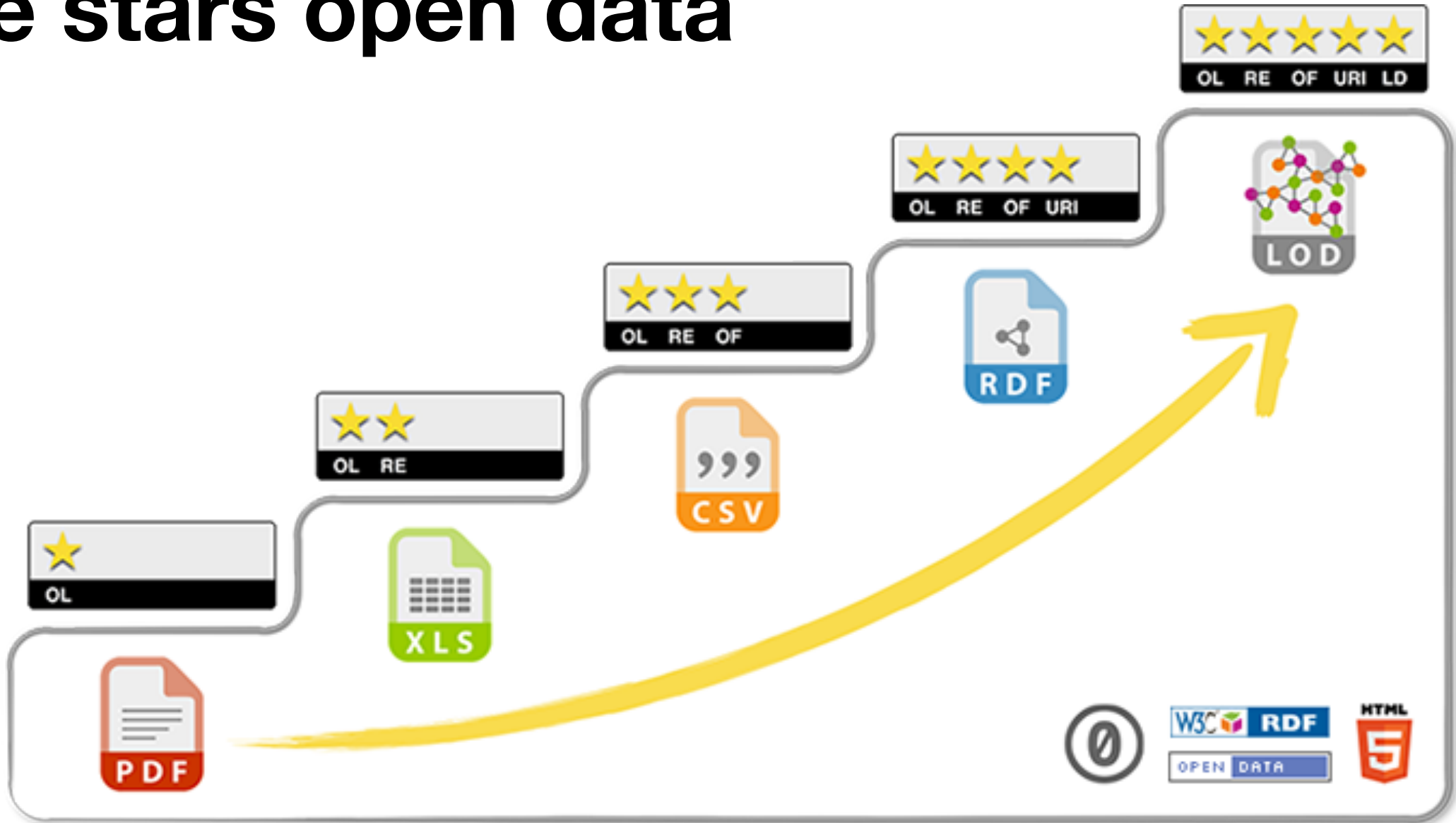
- representing data in machine-readable forms
- allowing links to be created

This is a pity, as data describe real objects with meanings and relationships between them

The open data is a flat, boring world devoid of meaning



Five stars open data

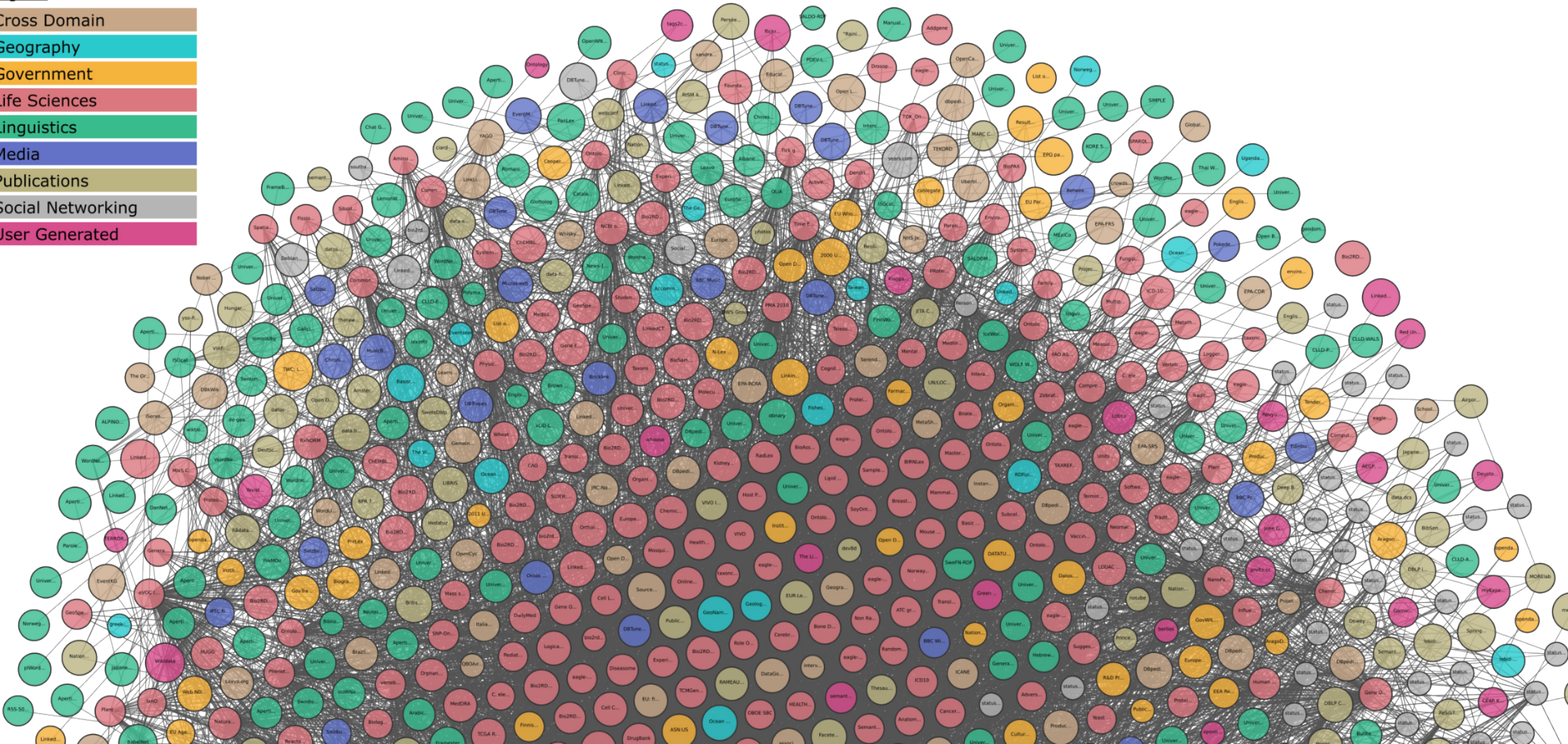


The Linked Open Data Cloud

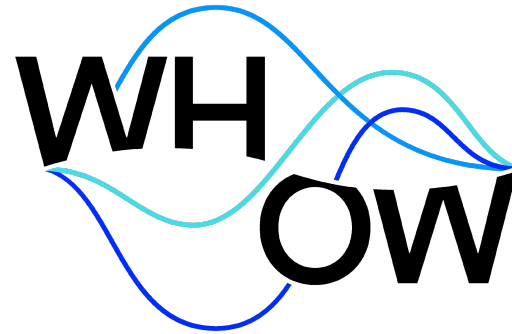
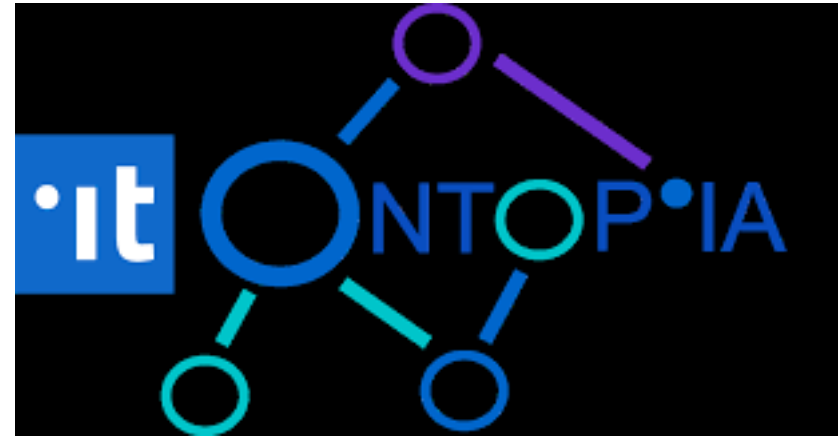
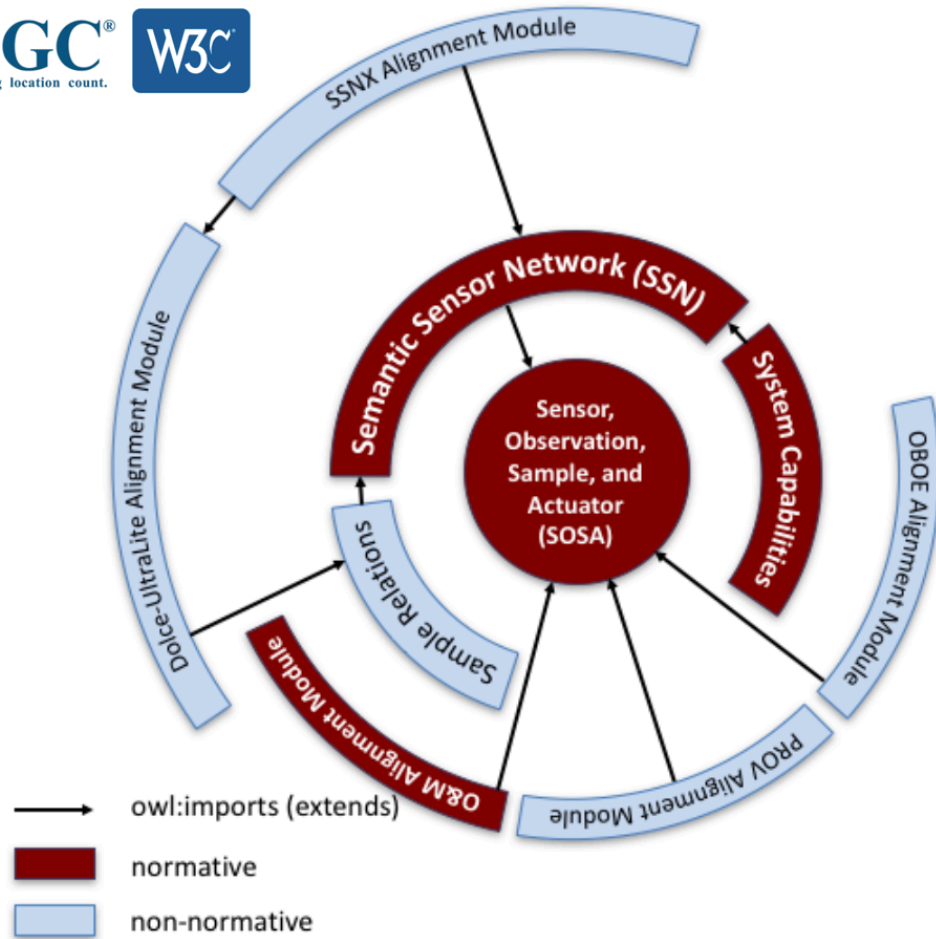


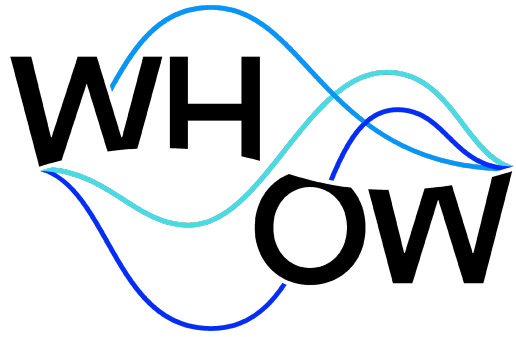
Legend

- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated

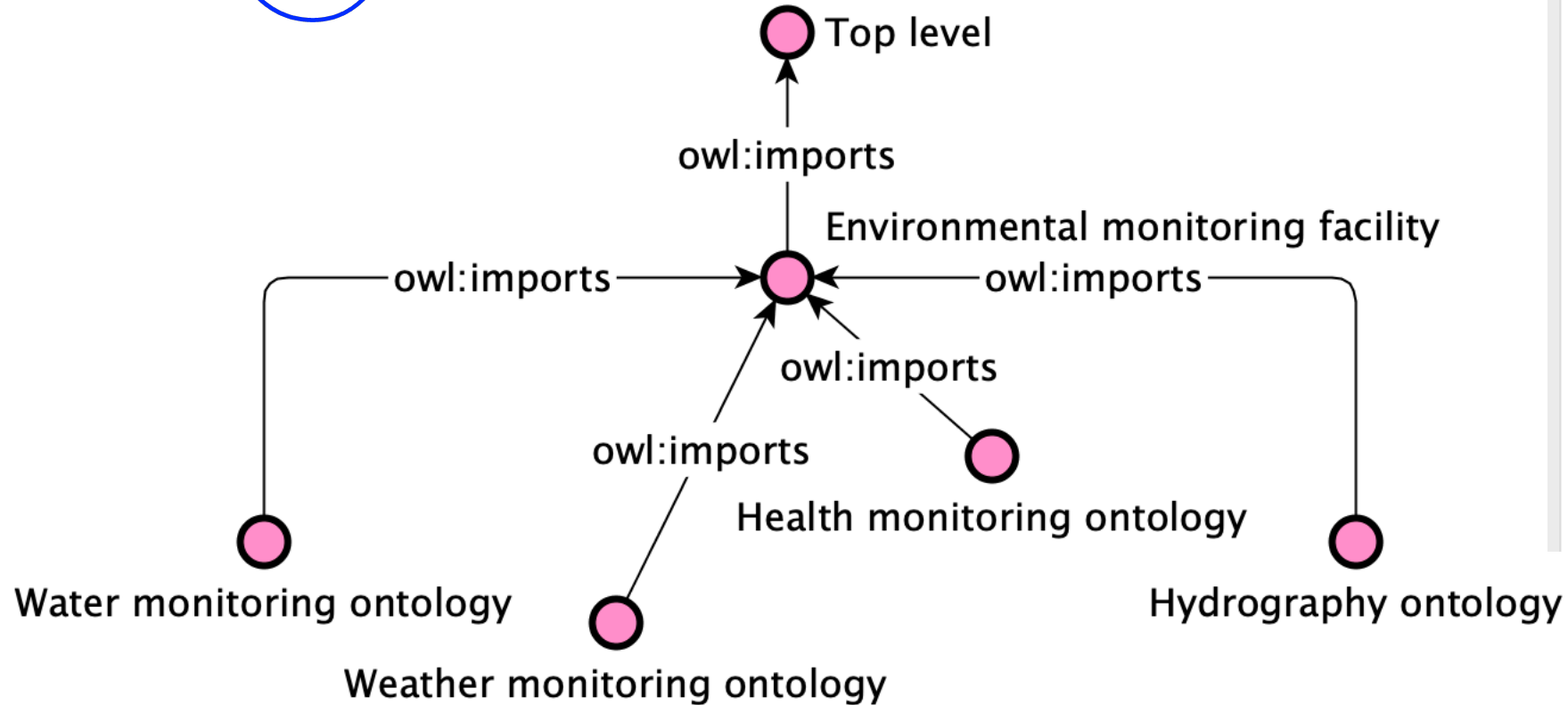


Existing ontology networks





ontology network

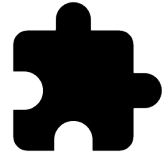


Ontology metrics:	
Metrics	
Axiom	4020
Logical axiom count	618
Declaration axioms count	492
Class count	201
Object property count	199
Data property count	62
Individual count	24
Annotation Property count	57
DL expressivity	SRIQ(D)
Class axioms	
SubClassOf	179
EquivalentClasses	5
DisjointClasses	24



WHOW toolkit

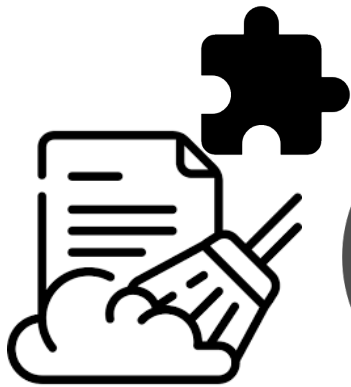
Workflow
manager



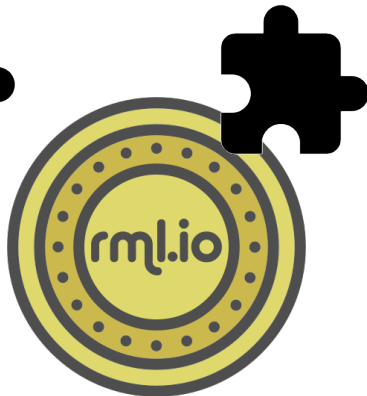
API
manager



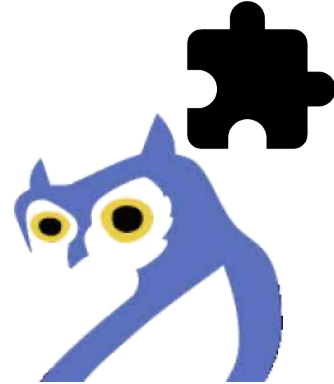
User
interfaces



Data
cleanser



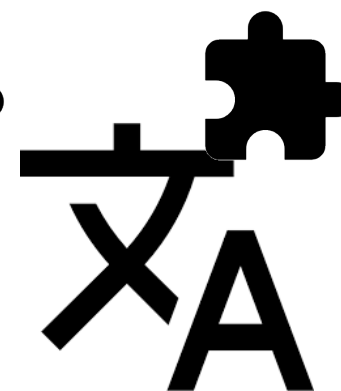
Triplifier



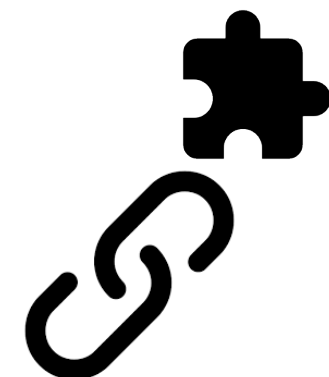
Hermit
Reasoner



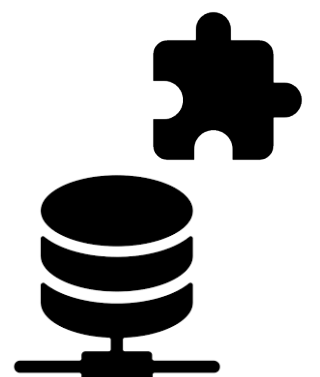
Data
Validator



eTranslation

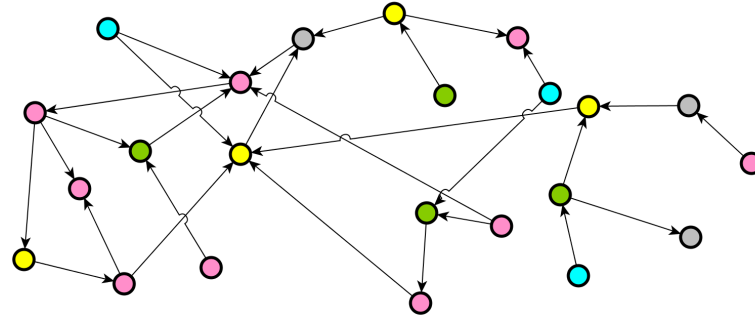


LIMES
Linking



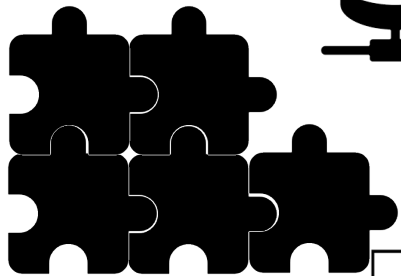
SPARQL
Triplestore

Customised and decentralised recipes for Knowledge Graph creation



Workflow manager

SPARQL



Cleanser

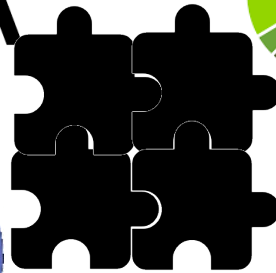
Triplifier

Validator

Provider A

eTranslation

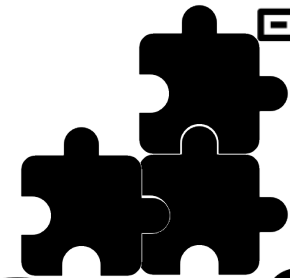
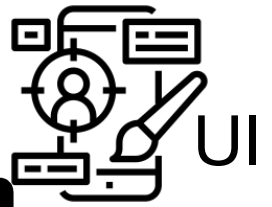
API



Reasoner

SPARQL

Provider B



Triplifier

Linking

Provider C

WHOW resources and key persons



Technical report



Ontology network



Toolkit source code



Giorgia Lodi



Anna Sofia Lippolis



Andrea Nuzzolese





 whowproject.eu
 info@whowproject.eu



Thank you!

Creative Commons Attribution 4.0
International (CC BY 4.0)



Co-financed by the Connecting Europe
Facility of the European Union

