RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

MINES Saint-Étienne | Institut Mines-Télécom

Semantic Interoperability for Smart Applications in the Energy Domain

Ontologies for the Internet of Things to Data Spaces applications

Maxime Lefrançois

Associate Professor
IMT - MINES Saint-Étienne - Institut Henri Fayol
Member of Laboratoire LIMOS - UMR CNRS 6158
https://maxime-lefrancois.info

ESIA 2025 – 2025-04-16

Outline

Semantic Interoperability for Smart Applications in the Energy Domain

- Backgroud on data interoperability
- Open Data and example
- Example of research based on open data: territorial energy planning
- Beyond open data: Data spaces
- Example of a large project on the European Data Space: Omega-X
- Background on knowledge graphs and ontologies
- Cross-domain interoperability with the ETSI Smart Application REFerence ontology
- The Omega-X Common Semantic Data Model
- Drive the point home with the Data Act

European Commission

EN

Search

Home > Strategy and policy > Priorities > A Europe fit for the digital age > European data strategy

European data strategy

Making the EU a role model for a society empowered by data

Access to data and the ability to use it are essential for innovation and growth. Data-driven innovation can bring major and concrete benefits, such as

- personalised medicine
- improved mobility
- better policymaking
- upgrading public services

1)

https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283

2)

https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy

Projected figures 2025

530%

increase of global data volume

From 33 zettabytes in 2018 to 175 zettabytes

€829 billion

value of data economy in the EU27

From €301 billion (2.4% of EU GDP) in 2018

10.9 million

data professionals in the EU27

From 5.7 million in 2018

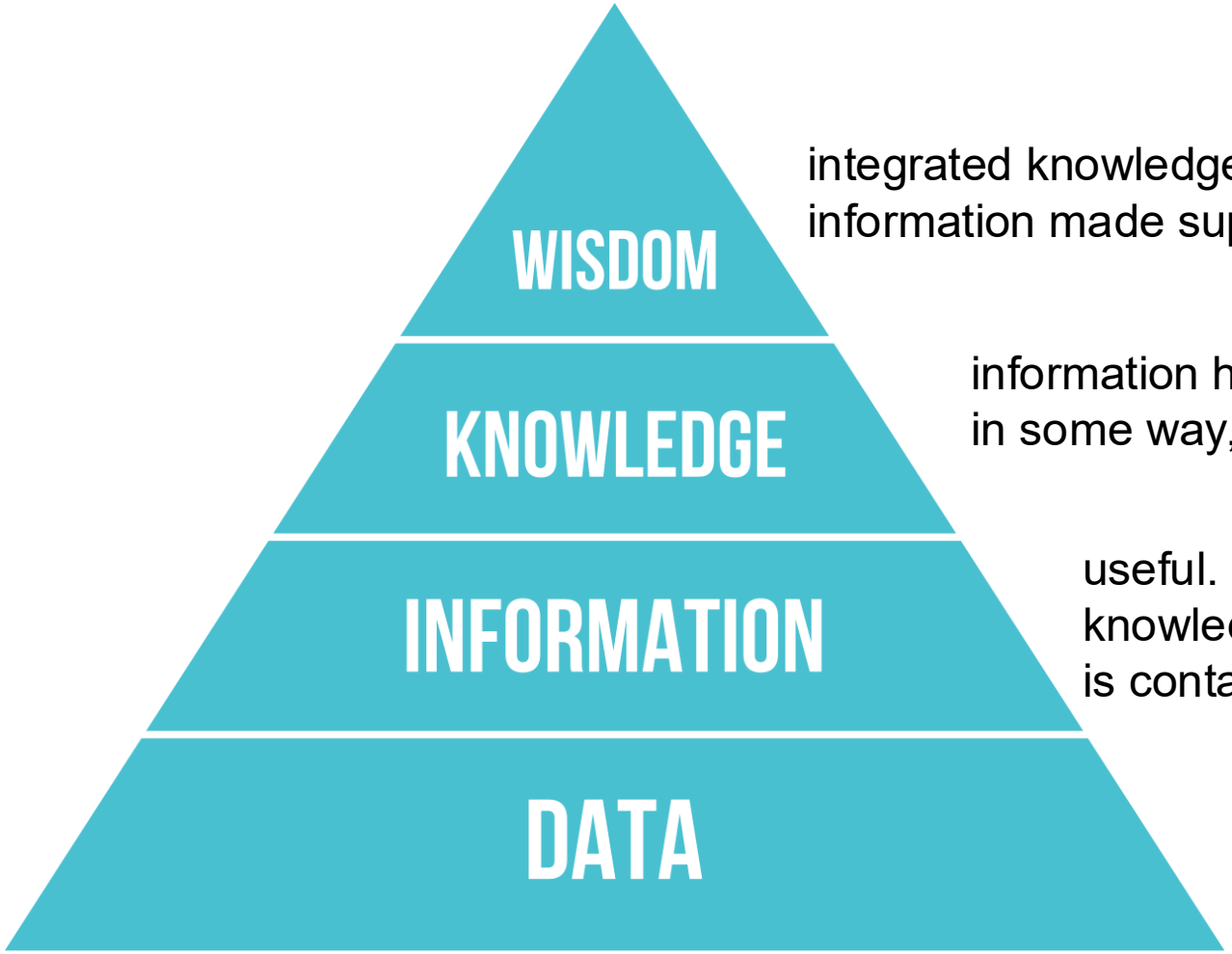
65%

Percentage of EU population with basic digital skills

From 57% in 2018

https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en#projected-figures-2025

The DIKW pyramid (~1982)



integrated knowledge—
information made super-useful

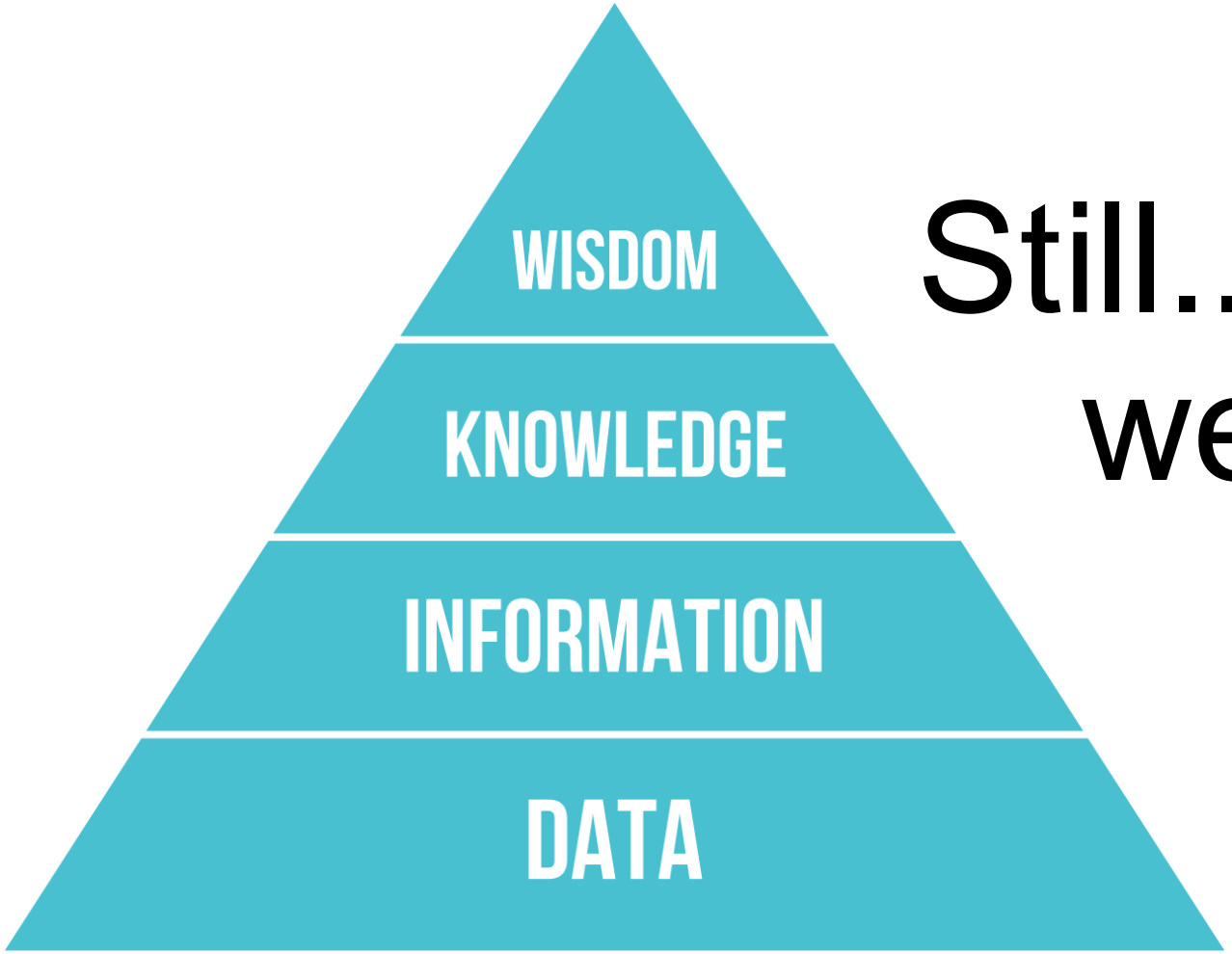
information having been processed, organized or structured in some way, or else as being applied or put into action

useful.
knowledge by description ("information is contained in descriptions")

useless.
symbols or signs, representing stimuli or signals, “of no use until... in a usable (that is, relevant) form”

https://en.wikipedia.org/wiki/DIKW_pyramid , and chosen definitions

The DIKW pyramid (~1982)



Still…

we often use

« Data »

to generalize

Definitions on Data

— ISO/IEC20546:2019 (Big data – Overview and vocabulary)

Data
Re-interpretable representation of information in a formalized manner suitable for communication, interpretation, or processing
Note 1 to entry: Data can be processed by humans or by automatic means.

Dataset
Identifiable collection of data available for access or download in one or more formats

Information
Data that are processed, organised and correlated to produce meaning.
Note 1 to entry: Information concerns facts, concepts, objects, events, ideas, processes, etc.

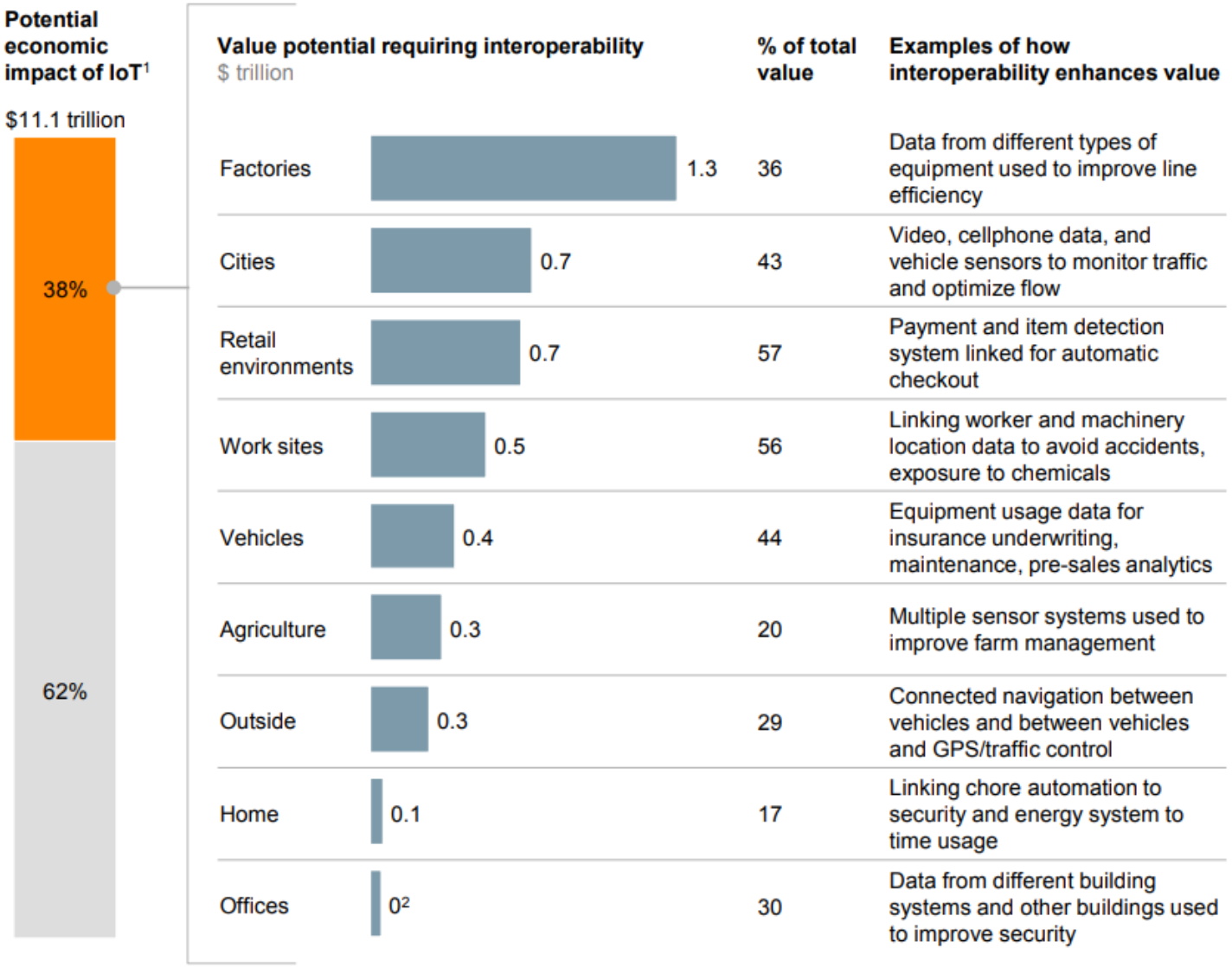
Metadata
Data about data or data elements, possibly including their data descriptions and data about data ownership, access paths, access rights and data volatility

Interoperability unlocks the value of data

Interoperability as an enabler for the potential value of data

ex:
Internet of Things

Nearly 40 percent of economic impact requires interoperability between IoT systems



1 Includes sized applications only; includes consumer surplus.
2 Less than $100 billion.
NOTE: Numbers may not sum due to rounding.

SOURCE: Expert interviews; McKinsey Global Institute analysis

https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world

Different facets of interoperability (ISO/IEC definitions)

Transport interoperability
interoperability where information exchange uses an established communication infrastructure between the participating systems

— *ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary*

Syntactic interoperability
interoperability such that the formats of the exchanged information can be understood by the participating systems

— *ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary*

Semantic interoperability
interoperability so that the meaning of the data model within the context of a subject area is understood by the participating systems

— *ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary*

Behavioural interoperability
interoperability so that the actual result of the exchange achieves the expected outcome

— *ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary*

Policy interoperability
interoperability while complying with the legal, organizational, and policy frameworks applicable to the participating systems

— *ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary*

Different facets of interoperability (ISO/IEC definitions)

Facets	Aim	Objects	Requirements	Examples
Transport	Data transfer between systems	Physical connections Signals	Protocols of data transfer	HTTP/S, MQTT
Syntactic	Receive data in an understood format	Data	Standardized data exchange formats	JSON, XML, ASN.1
Semantic	Receive data using an understood data information model	Programmatic interface	Common interpretation of data information model	Directories, data keys, ontologies
Behavioural	Obtain expected outcomes to interface operations	Information	Behavioural model(s) of the invoked IoT entity	UML models, pre- and post-conditions, constraint specifications
Policy	Assurance that interoperating systems follow applicable regulatory and organizational policies	Regulatory and organizational polices and interoperation context	Conditions and control for use and access	Security policies of IoT system stakeholders, restriction on cross-border data transfer, regulations controlling PII

Source: ISO/IEC 19941, *Information technology – Cloud computing – Interoperability and portability*

Data Interoperability

Data interoperability addresses the ability of systems and services that create, exchange and consume data to have clear, shared expectations for the contents, context and meaning of that data.

— https://datainteroperability.org/

Syntactic interoperability

interoperability such that the formats of the exchanged information can be understood by the participating systems

— ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary

Semantic interoperability

interoperability so that the meaning of the data model within the context of a subject area is understood by the participating systems

— ISO/IEC 22123-1:2021, Cloud computing — Part 4: Vocabulary

Demo: data heterogeneities

https://samples.openweathermap.org/data/2.5/weather?id=2172797&appid=b6907d289e10d714a6e88b30761fae22

and

https://www.prevision-meteo.ch/services/json/lausanne

Recap: data heterogeneities

Different modeling choices were made, which make these two services completely non-interoperable:

- the lat/long coordinates: string vs number
- the UNIX timestamps vs dates and times
- the choice of keys and the semantics (meaning) of the values
- the units of temperature, pressure, wind speed, …
- the semantics of wind direction
- the value for "icon": "03n" - (if we follow our nose on the website, we may figure out it refers to http://openweathermap.org/img/w/03n.png)
- the country codes ISO 3166-1 ALPHA-2 and ISO 3166-1 ALPHA-3 (example of Australia and Austria)

Standard Data Models for the Energy Domain ?

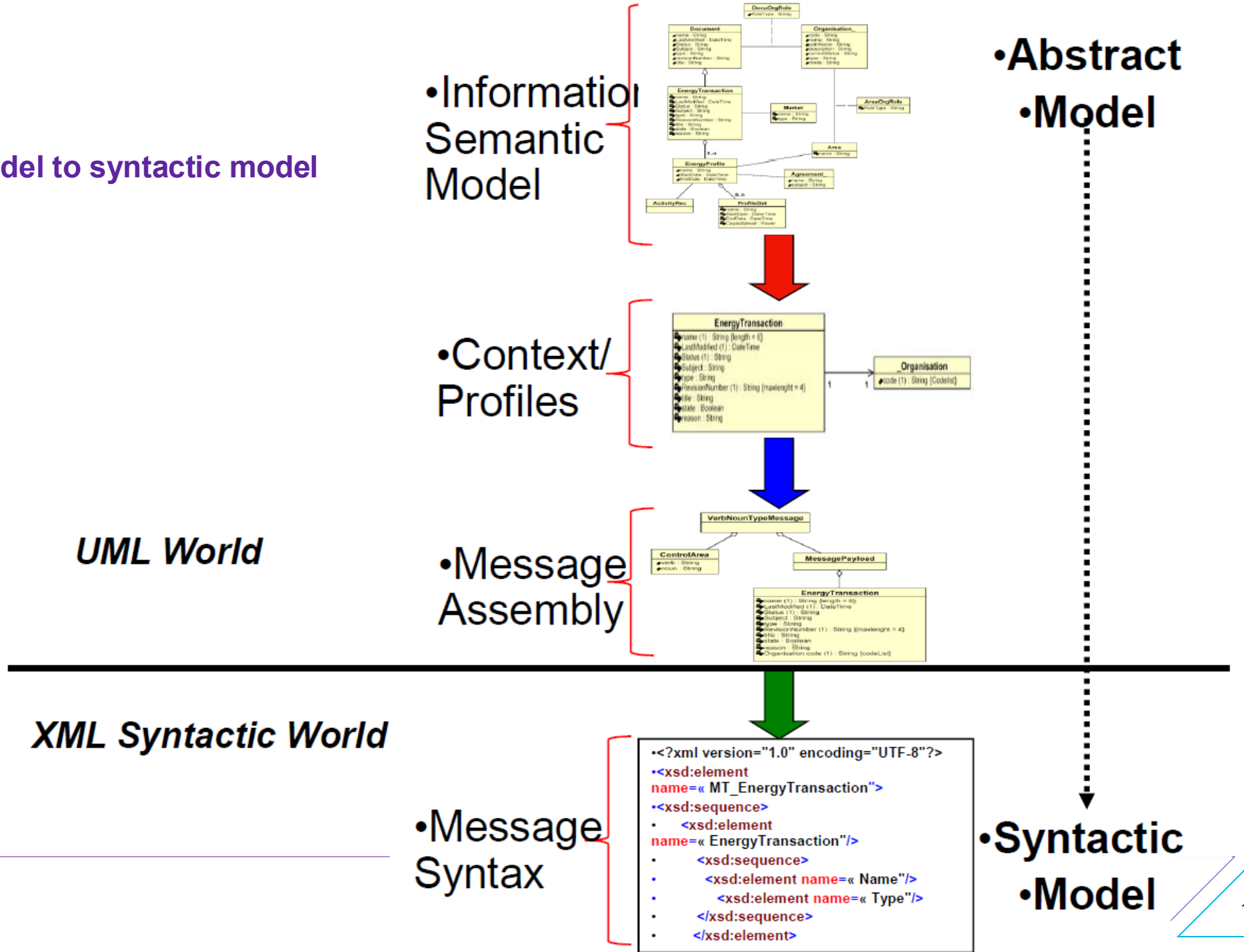
Standard Data Models for the Energy Domain ?

- We can go with a list:
  - IEEE 2030.2 : communication in transport and distribution networks
  - IEEE C37.118: synchrophasers (measures phase shift @ 30Hz = real time map)
  - IEC 61850: substation automation
  - IEC CIM: Common Information Model: general description of the energy grid
  - OpenADR: Automated Demand/Response
  - KNX, Bac.net, …
  - Zigbee, Matter, LoRa,
  - …
- **Too broad ! Define scope of "*Energy Domain*".**
  - Is it the network ? Substations ? Metering ? Flexibility ? Local Energy Communities ? Renewables ? Electric Vehicles ?

Interoperability

CIM

From information model to syntactic model



Open Data generates societal value

Open Data in the US

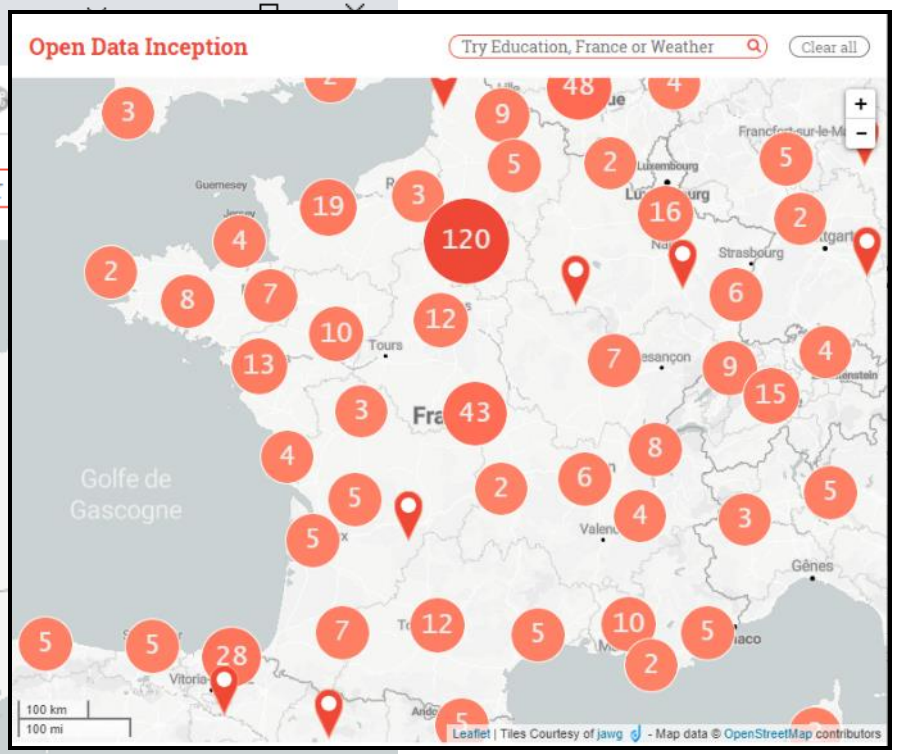
data.gov

Type of site	Government Web site
Available in	English
Owner	Government of the United States
URL	data.gov
Commercial	No
Registration	Optional
Launched	May 30, 2009; 13 years ago
Current status	Active

- Data.gov 2009
- Legal framework:
  - The U.S. Open Government Directive of December 8, 2009, required that all agencies post **at least three high-value data sets online** and register them on Data.gov **within 45 days**
  - OPEN Government Data Act, as part of the Foundations for Evidence Based Policymaking Act (2019)

Open Data in France

- France at the forefront of Open Data in Europe:
  - Légifrance 1999
- Legal framework:
  - "The society has the right of requesting account from any public agent of its administration." (Declaration of rights of man and of the citizen of 1789)
  - Law on the liberty of access to administrative documents (1978)
  - Euopean directove 2003 + French Law 2005 + Decree 2011
  - Bill on a Digital Republic (2016)
  - The law on Energy Transition (2015)
- 2014: Chief Data Officer in the French public administration



https://opendatainception.io/

View as list on Opendatasoft's website **Add a portal** Download the latest version here View terms of use

Open Data in Europe

EUROPEAN
DATA PORTAL

The European Data Portal:
Opening up Europe's public data

data.europa.eu/europeandataportal

data.europa.eu
The official portal for European data

Economic impact of Open Data

Open data market size

- €184.45 billion open data market size in 2019
- €199.51 - €334.20 billion open data market size forecast for 2025

Open data employment

- 1.09 million open data employees in 2019
- 1.12 – 1.97 million open data employees forecast for 2025

Open data potential per sector

- 15.7% growth expected from high impact and high potential sectors
- High impact:
- High potential:

For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

Efficiency gains

- Saving lives, e.g. 54 - 202 thousand lives saved by faster emergency response
- Saving time, e.g. 27 million hours saved in public transport
- Saving the environment, e.g. 5.8 Mtoe* saved by reducing household energy consumption
- Improving language services with open data, e.g. by increasing machine translation

Cost savings

- Saving healthcare costs, e.g. €312 - €400 thousand due to faster first aid by bystanders
- Saving labour costs, e.g. €13.7 - €20 billion by reducing time spent in traffic
- Saving costs on energy bills, e.g. €79.6 billion due to more solar energy production
- Saving public sector costs, e.g. €1.1 billion by lower translation costs

Open data in organisations

- 49% of data used by surveyed organisations is open data and 77% of organisations plan to use more data
- 46% of organisations' revenues are impacted by open data and 73% of organisations expect the impact to increase
- 70% of surveyed organisations create data internally, of which 58% publish some of it as open data

* Million tonnes of oil equivalent

For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

https://data.europa.eu/sites/default/files/the-economic-impact-of-open-data.pdf (2020)

https://data.europa.eu/sites/default/files/the-economic-impact-of-open-data.pdf

EUROPEAN DATA PORTAL

Re-using Open Data

A study on companies transforming Open Data into economic & societal value

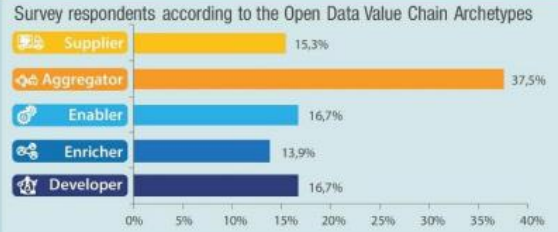
EUROPEAN DATA PORTAL

TOP 3 SOURCES OF REVENUE OF OPEN DATA COMPANIES

- 21% Selling products & services
- 42% Selling services
- 10% Selling products

OPEN DATA ARCHETYPES

Survey respondents according to the Open Data Value Chain Archetypes



TOP 3 PROFILES OPEN DATA COMPANIES ARE LOOKING FOR

ESSENTIAL CONDITIONS FOR RE-USING OPEN DATA

- High quality Open Data
- Systematic and continued publication of Open Data

?

76% OF THE ORGANISATIONS USING OPEN DATA FORESEE TO RECRUIT NEW EMPLOYEES

TAKING OPEN DATA TO THE NEXT LEVEL

- Design your Open Data provision strategy based on user demands
- Standardise and harmonise your Open Data
- Share your story on the use of Open Data

https://data.europa.eu/sites/default/files/re-using_open_data.pdf (2020)

EUROPEAN DATA PORTAL

Re-using Open Data

A study on companies transforming Open Data into economic & societal value

https://data.europa.eu/sites/default/files/re-using_open_data.pdf (2020)

EUROPEAN DATA PORTAL

Top 3 most used Open Data domains

27.3% Statistical

25.8% Geospatial

19.5% Companies

Most often combined Open Data categories

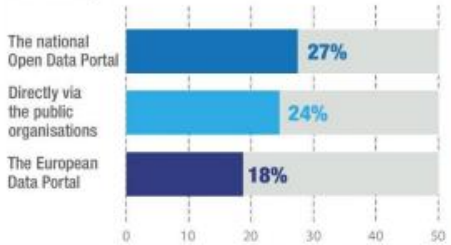
Other Open Data categories used

In clockwise order:

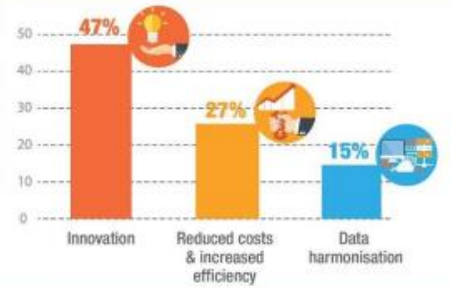
- Agriculture, fisheries, Forestry & Foods
- Energy
- Economy & Finance
- International Issues
- Government & Public Sector
- Justice, Legal System & Public Safety
- Education, Culture & Sport
- Health
- Science & Technology

Top 3 platforms to access Open Data

Provided by



Main benefits of working with Open Data



European Commission

Home > Strategy > Priorities 2019-2024 > A Europe fit for the digital age > European data strategy

European legislation on open data

(adopted 05/2022 – applicable 09/2023

The Directive on open data and the re-use of public sector information provides common rules for a European market for government-held data.

The **"Open Data Directive"** (EU) 2019/1024 entered into force on 16 July 2019

It replaced the **Public Sector Information (PSI) Directive** of 2003.

EU countries had to transpose Directive (EU) 2019/1024 by 16 July 2021.

The **Commission Implementing Regulation (EU) 2023/138** adopted a list of specific high-value datasets by way of an implementing act.

1)

https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data

European Commission

Home > Strategy > Priorities 2019-2024 > A Europe fit for the digital age > European data strategy

European legislation on open data

(adopted 05/2022 – applicable 09/2023

20.1.2023 EN Official Journal of the European Union L 19/43

COMMISSION IMPLEMENTING REGULATION (EU) 2023/138

of 21 December 2022

laying down a list of specific high-value datasets and the arrangements for their publication and re-use

(Text with EEA relevance)

2)

The **Commission Implementing Regulation (EU) 2023/138** adopted a list of specific high-value datasets by way of an implementing act.

http://data.europa.eu/eli/reg_impl/2023/138/oj

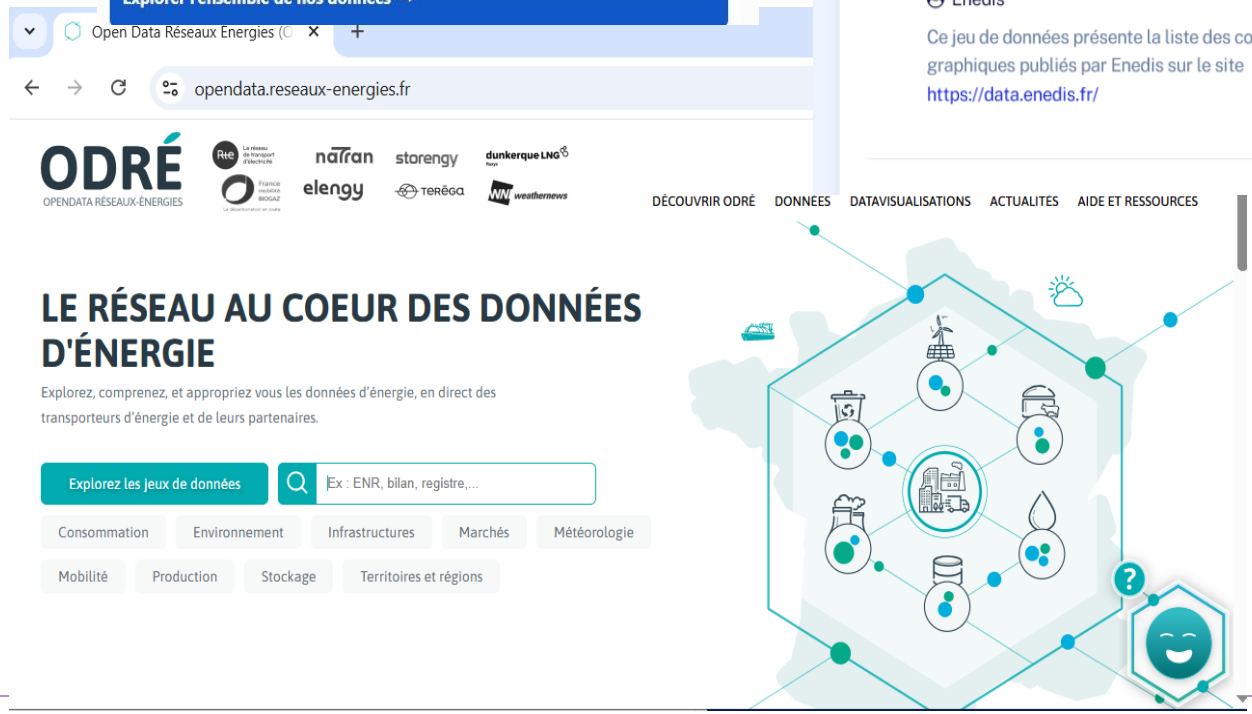
Actual impact (FR)

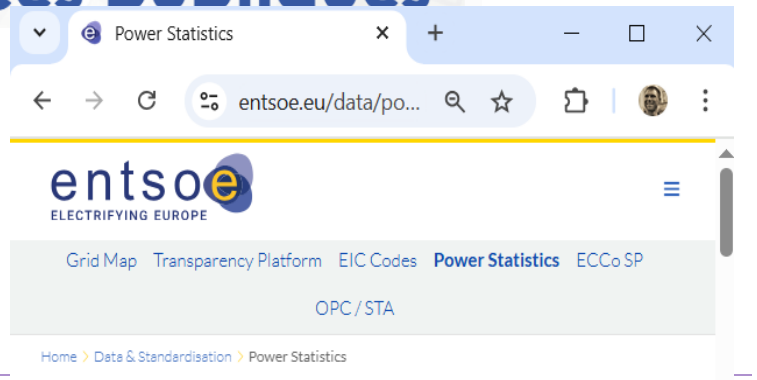
https://www.etalab.gouv.fr/

https://meteo.data.gouv.fr/

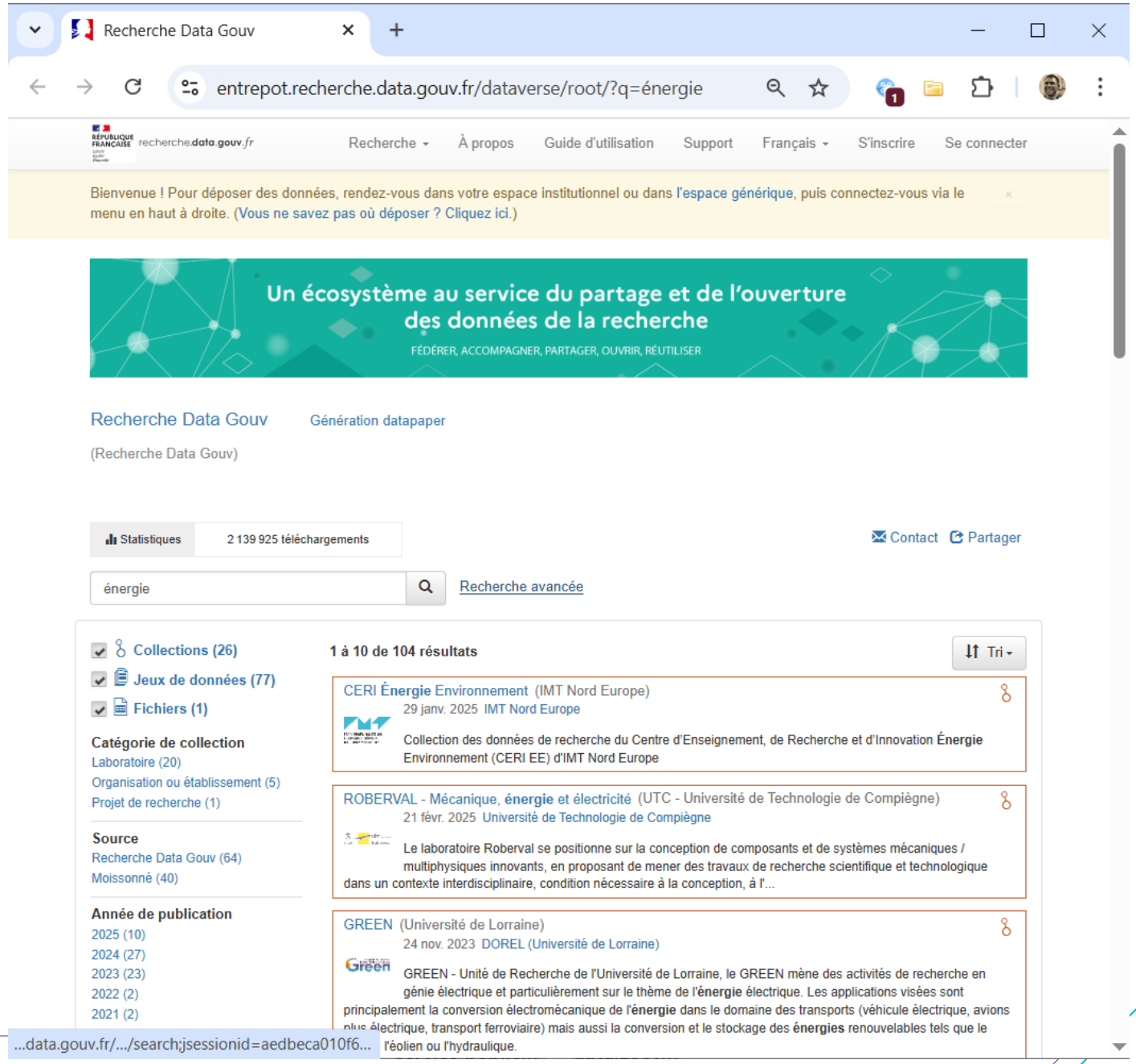
https://transport.data.gouv.fr/

Open Energy Data



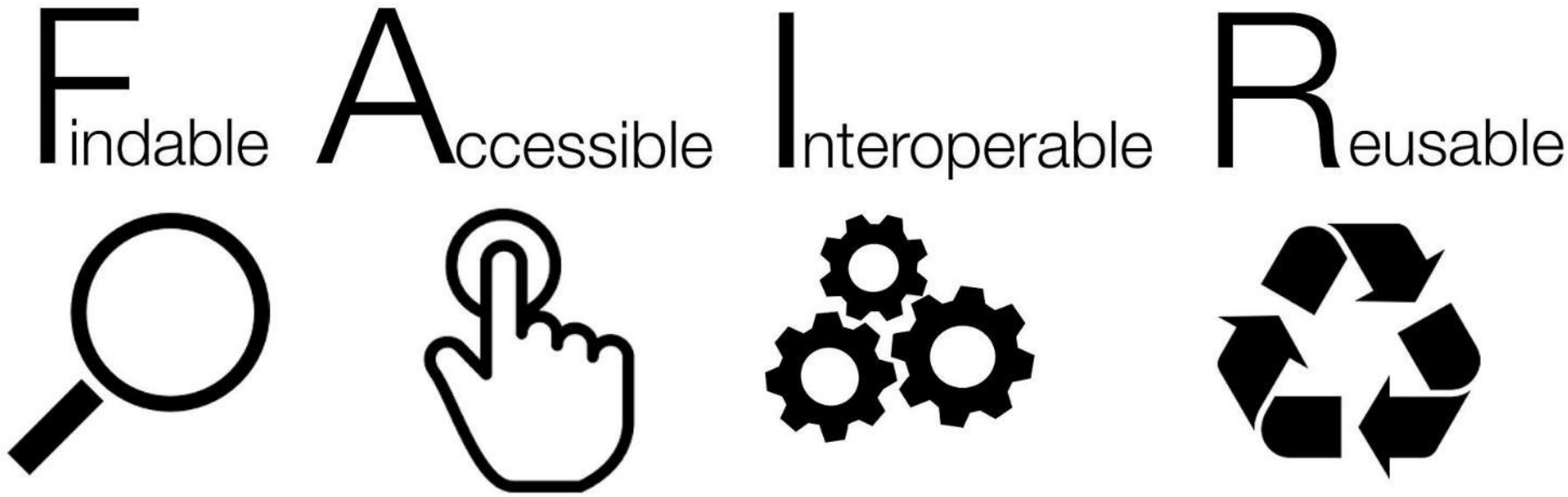


Open Research Data



Machine-actionability of data increases its value

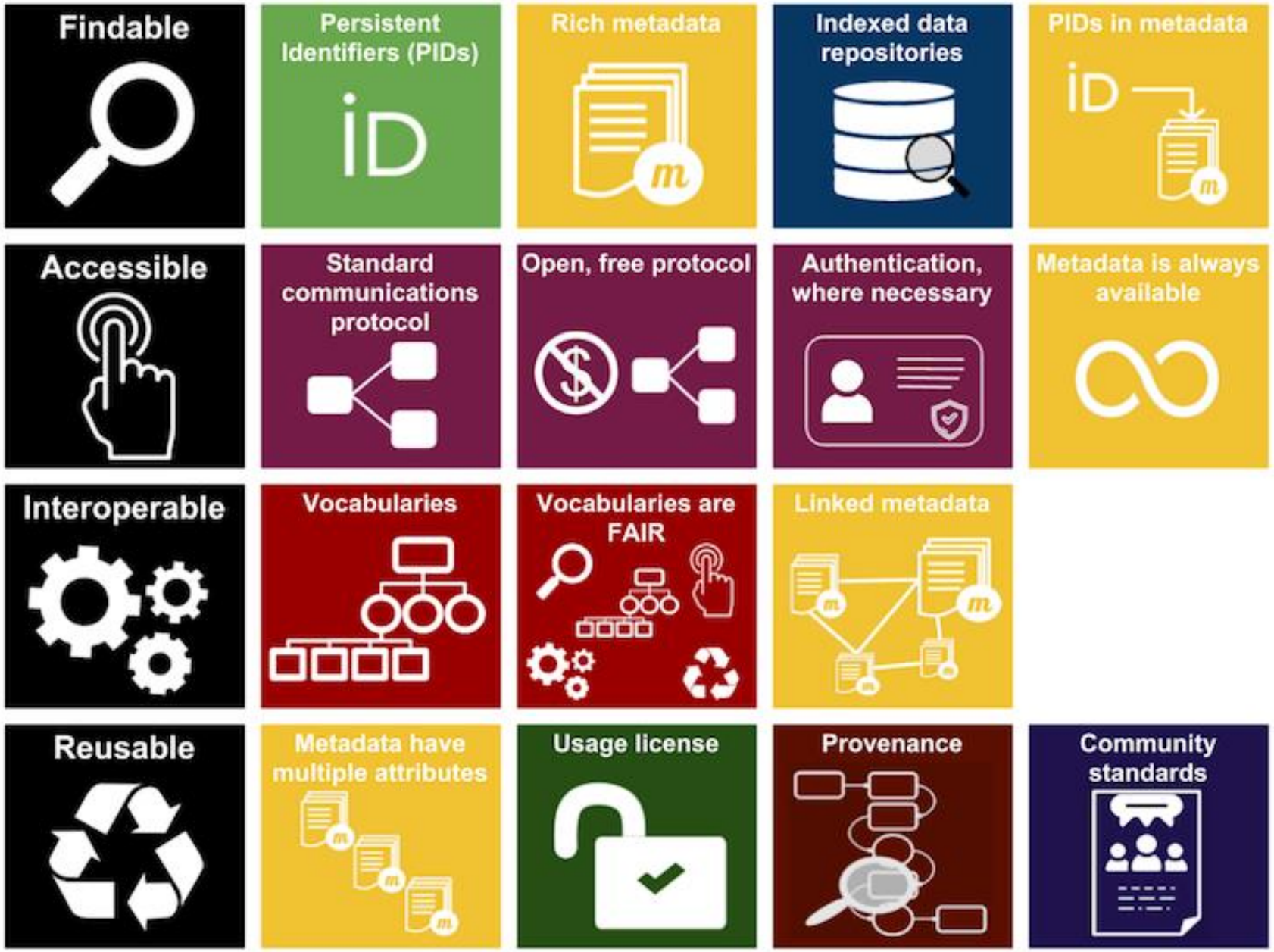
FAIR Principles



The FAIR principles emphasize machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.

https://www.go-fair.org/fair-principles/

FAIR Guiding Principles for scientific data management and stewardship (2016)



Source : Australian National Data Service (ANDS)

https://www.go-fair.org/fair-principles/

FAIR Guiding Principles for scientific data management and stewardship (2016)

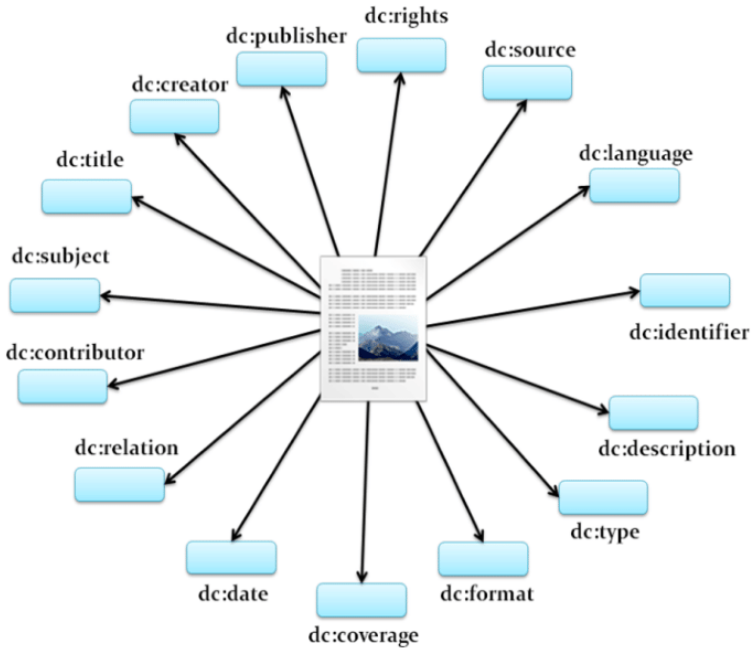
Metadata

"data that provides information about other data"

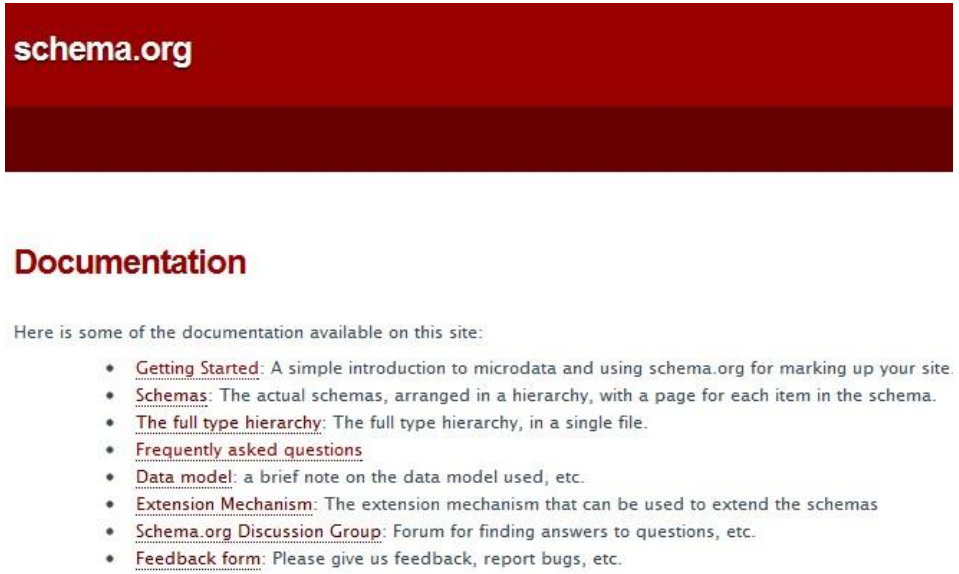
- **Descriptive metadata** – the descriptive information about a resource.
  - For discovery and identification.
  - Ex: title, abstract, author, keywords.
- **Structural metadata** – containers of data, how compound objects are put together
  - Ex, how pages are ordered to form chapters.
  - Ex: types, versions, relationships, and other characteristics of digital materials.
- **Administrative metadata** – the information to help manage a resource
  - Ex: resource type, permissions, time, when and how it was created.
- **Reference metadata** – contents and quality
  - For quality assessment of the data
  - Ex: conceptual metadata, quality metadata, methodological metadata.
- **Statistical metadata**, also called process data,
  - May describe processes that collect, process, or produce statistical data.
  - Number of rows, columns, etc.
- **Legal metadata** –
  - Ex: license, creator, copyright holder

https://en.wikipedia.org/wiki/Metadata

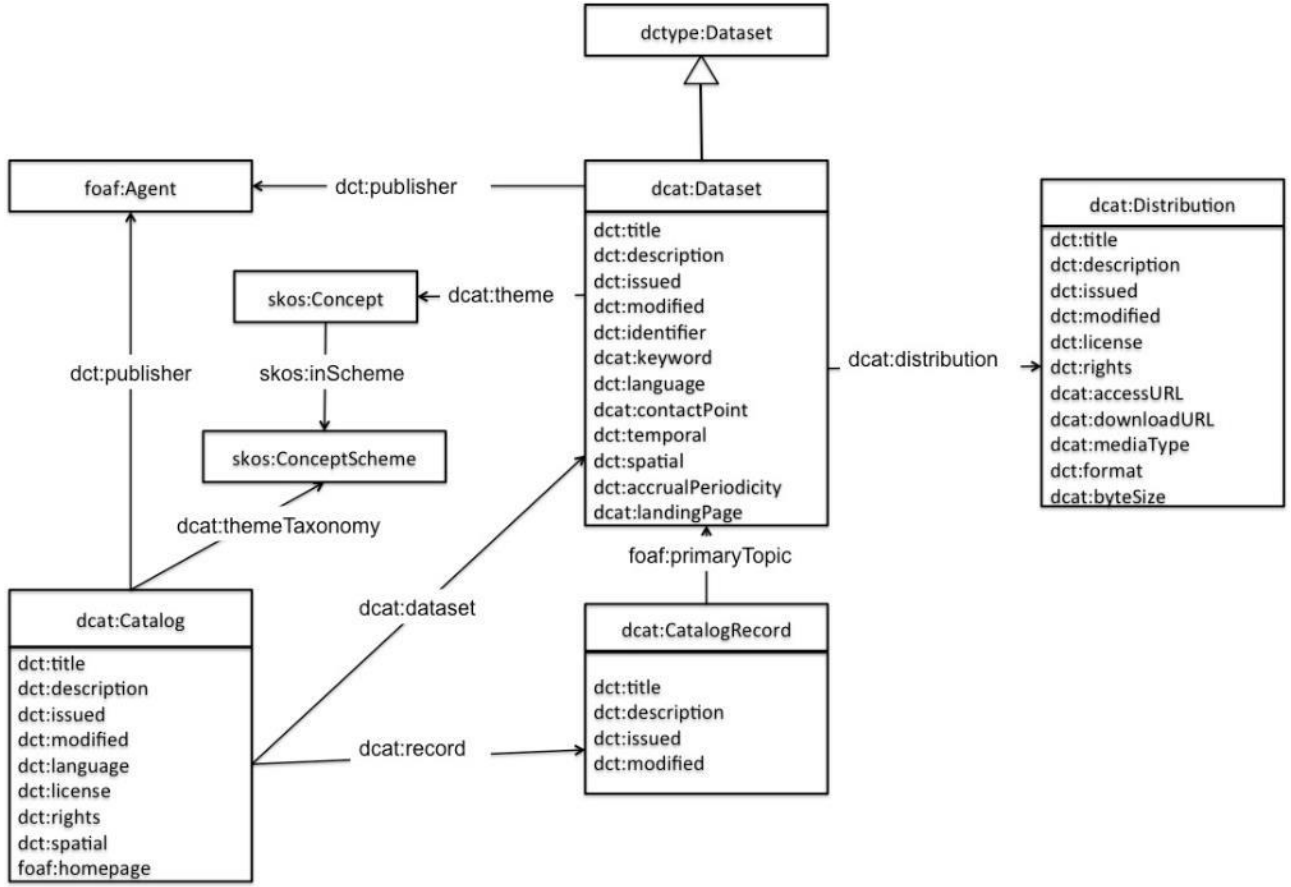
Example of standards for Metadata



Dublin Core



schema.org



Data Catalog Vocabulary (DCAT) - Version 2
W3C Recommendation 04 February 2020

Example of research based on the use of Open Data

Dual data/model approach for Territorial Energy Planning

Industrial PhD
Coline Baraize, Akajoule,
IMT Atlanque: Bruno Lacarriere, Pierrick Haurant
Mines Saint-Étienne: Maxime Lefrançois

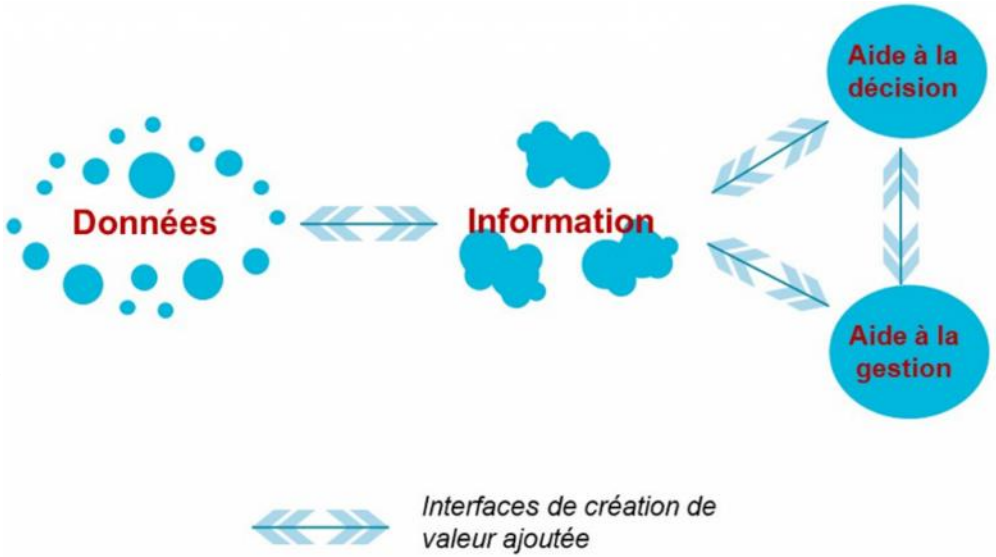
Context

Chair ValaDoE aims to improve heterogenous data used in energy planning

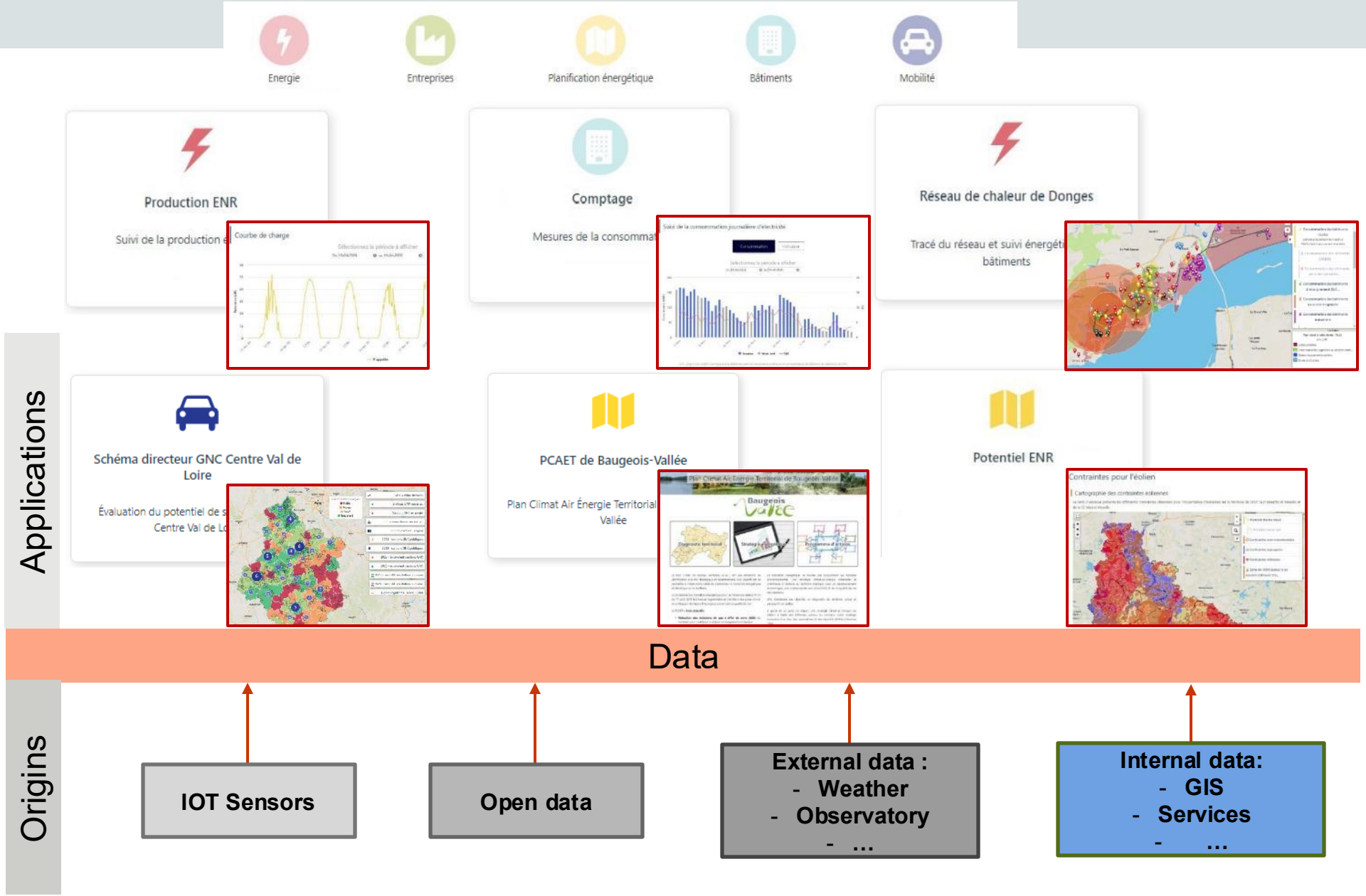
IMT Atlantique, Process Engineering

- Modeling-Optimization-Simulation
- Energy systems
- Territory energy data processing

Akajoule, consulting and engineering company in energy efficiency and renewable energy



Available Energy data of a territory



datajoule

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Territorial energy modeling

Territory

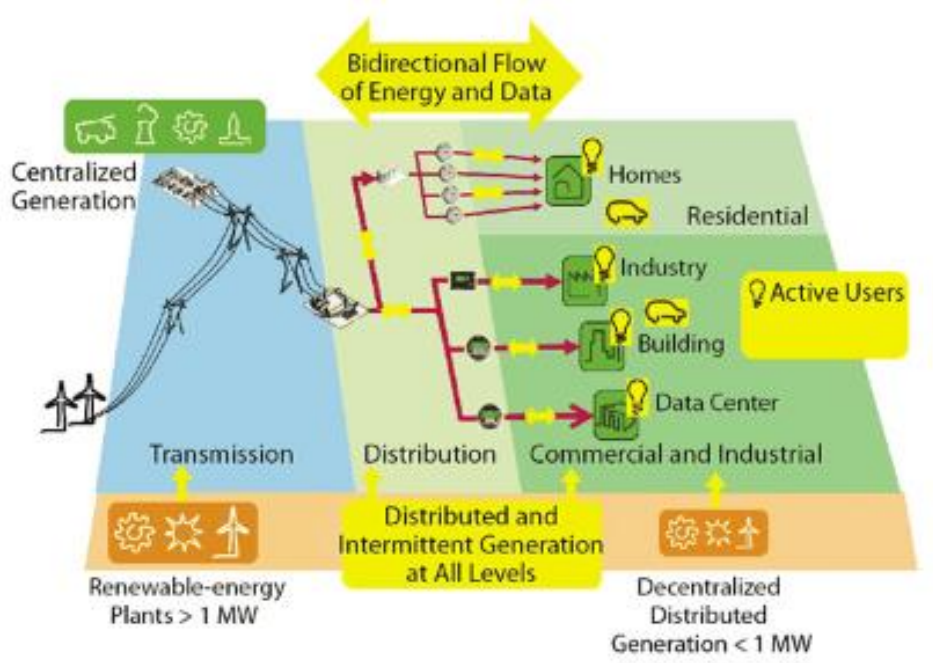


Figure 1: Prosumer Oriented Smart Grid[2]

Gillani, S., Laforest, F., & Picard, G. (2014, July). A Generic Ontology for Prosumer-Oriented Smart Grid. In EDBT/ICDT Workshops (Vol. 1133, pp. 134-139).

System modelling

Inputs → Energy system → Outputs

Existing system modelling

Irradiance → PV → Electricity

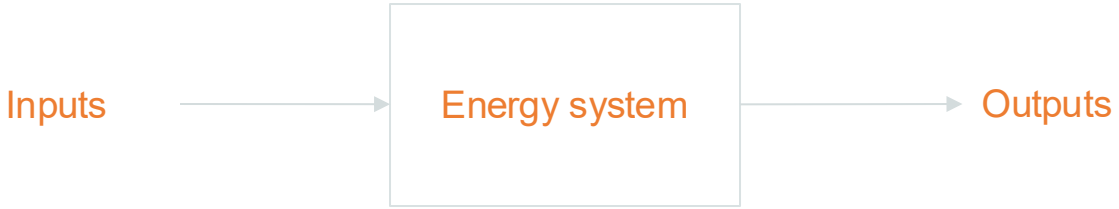
System modelling tested

Thermal energy → District heating network → Thermal energy

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

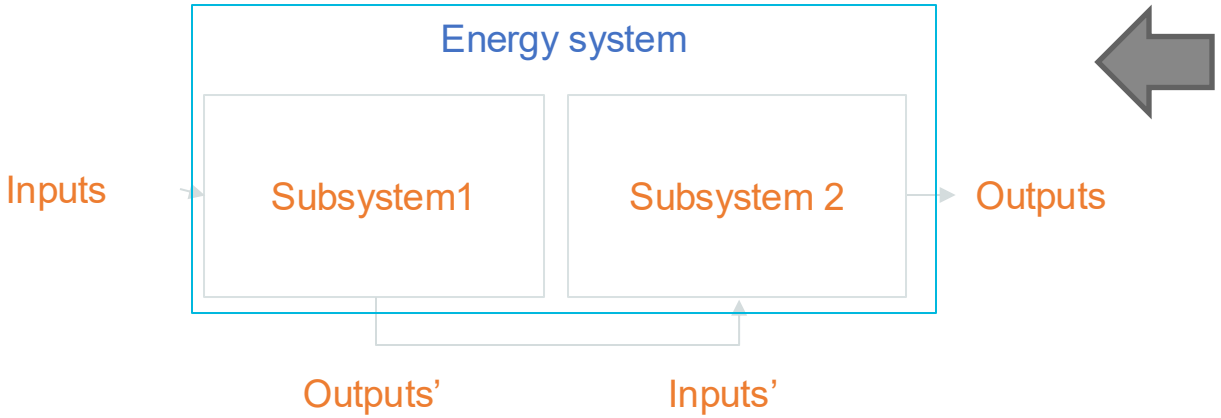
Territorial energy modeling

Maths modelling



$$Output = Efficiency * Input$$

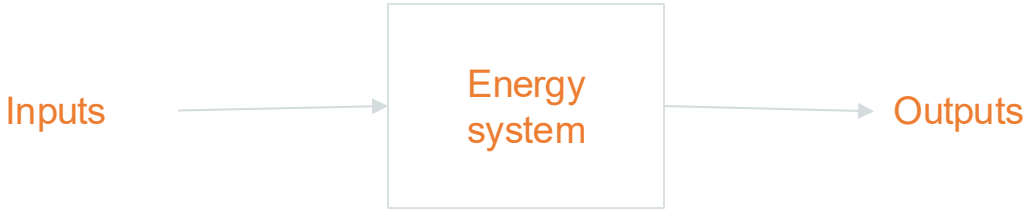
$$Ex: \boldsymbol{Efficiency(t) = a.OutdoorTemperature(t) + b}$$



$$Output = (Efficiency_2 * Efficiency_1) * Input$$

$$Efficiency = Efficiency_2 * Efficiency_1$$

System modelling



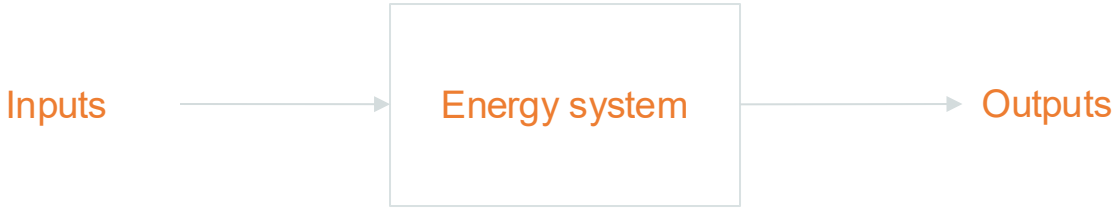
Existing system modelling

System modelling tested

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

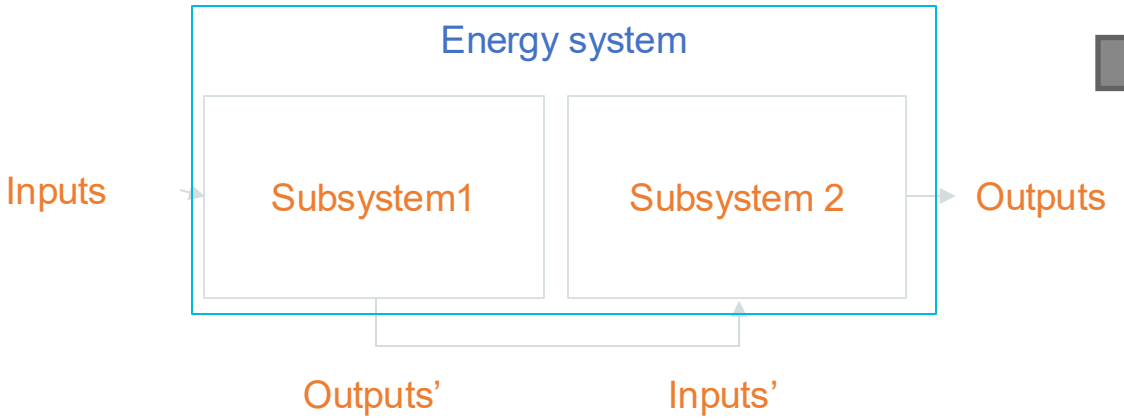
Territorial energy modeling

Maths modelling



$$Output = Efficiency * Input$$

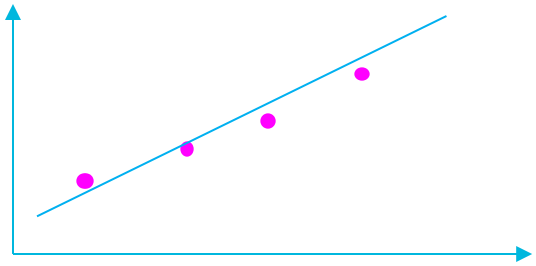
$$Ex: \boldsymbol{Efficiency(t) = a.OutdoorTemperature(t) + b}$$



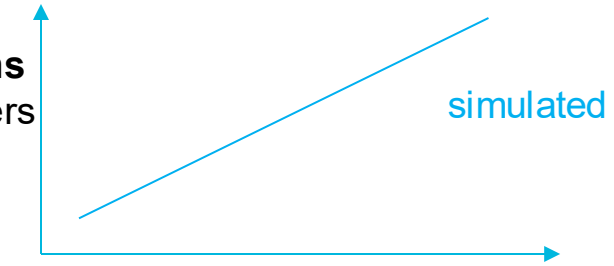
$$Output = (Efficiency_2 * Efficiency_1) * Input$$

$$Efficiency = Efficiency_2 * Efficiency_1$$

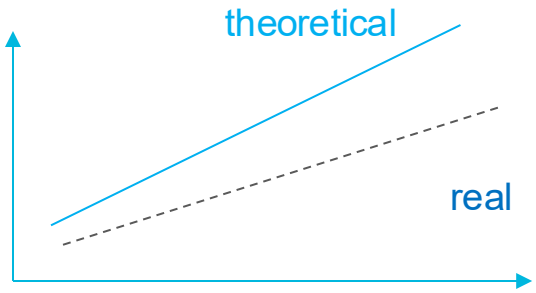
1) Learning or data analysis
→ Identify trends



2) Forecasts or simulations
→ Utility trends or parameters



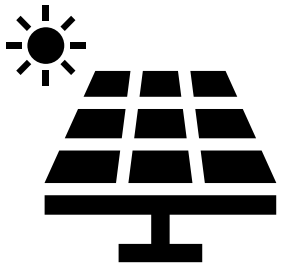
3) Validation
→ Identify errors or outliers



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Duality data/models ?

Energy system



Simplify process

Modelset

$$Output1 = Efficiency * Input$$
$$Efficiency(t) = a.OutdoorTemperature(t) + b$$
$$Efficiency(t) = c.InsideTemperature(t) + d$$
$$Output2 = Efficiency * Input$$
$$Efficiency(t) = e.InsideTemperature(t) + f$$

Add sensors

Aggregate (sum, mean, group by, filter, etc.)

Output

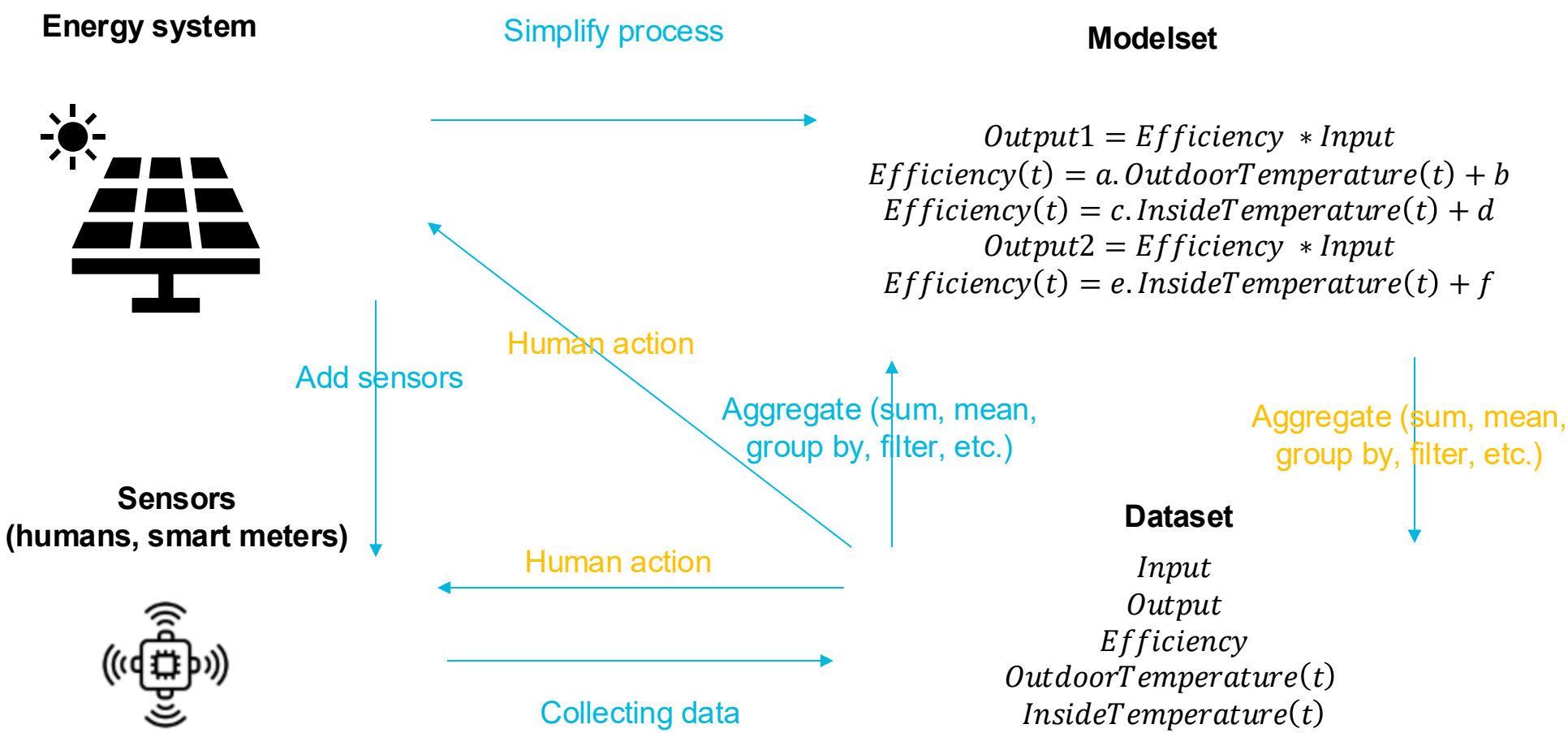
Sensors (humans, smart meters)

Collecting data

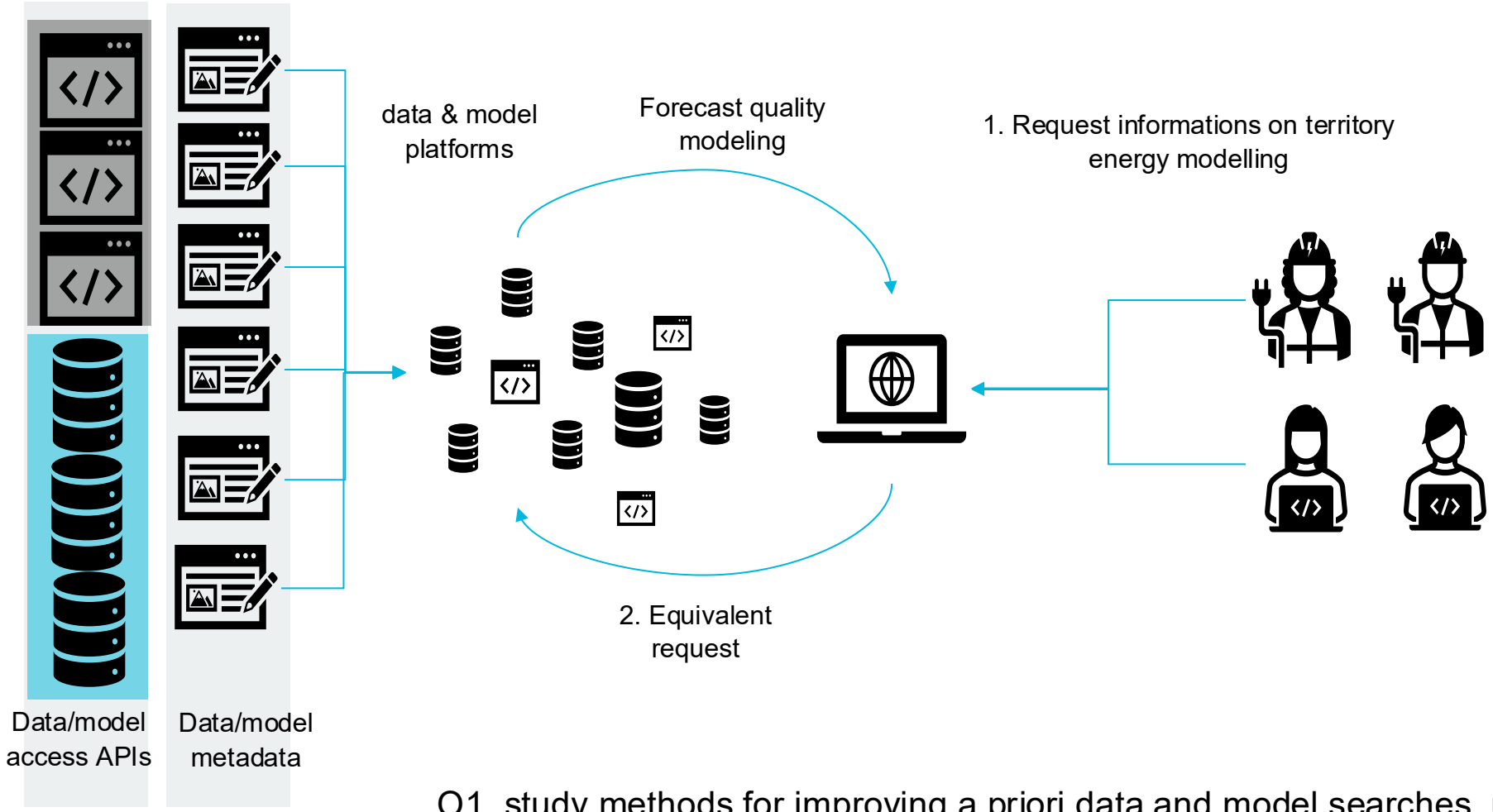
Dataset

$Input$
$Output$
$Efficiency$
$OutdoorTemperature(t)$
$InsideTemperature(t)$

Duality data/models ?



Thesis Objectives

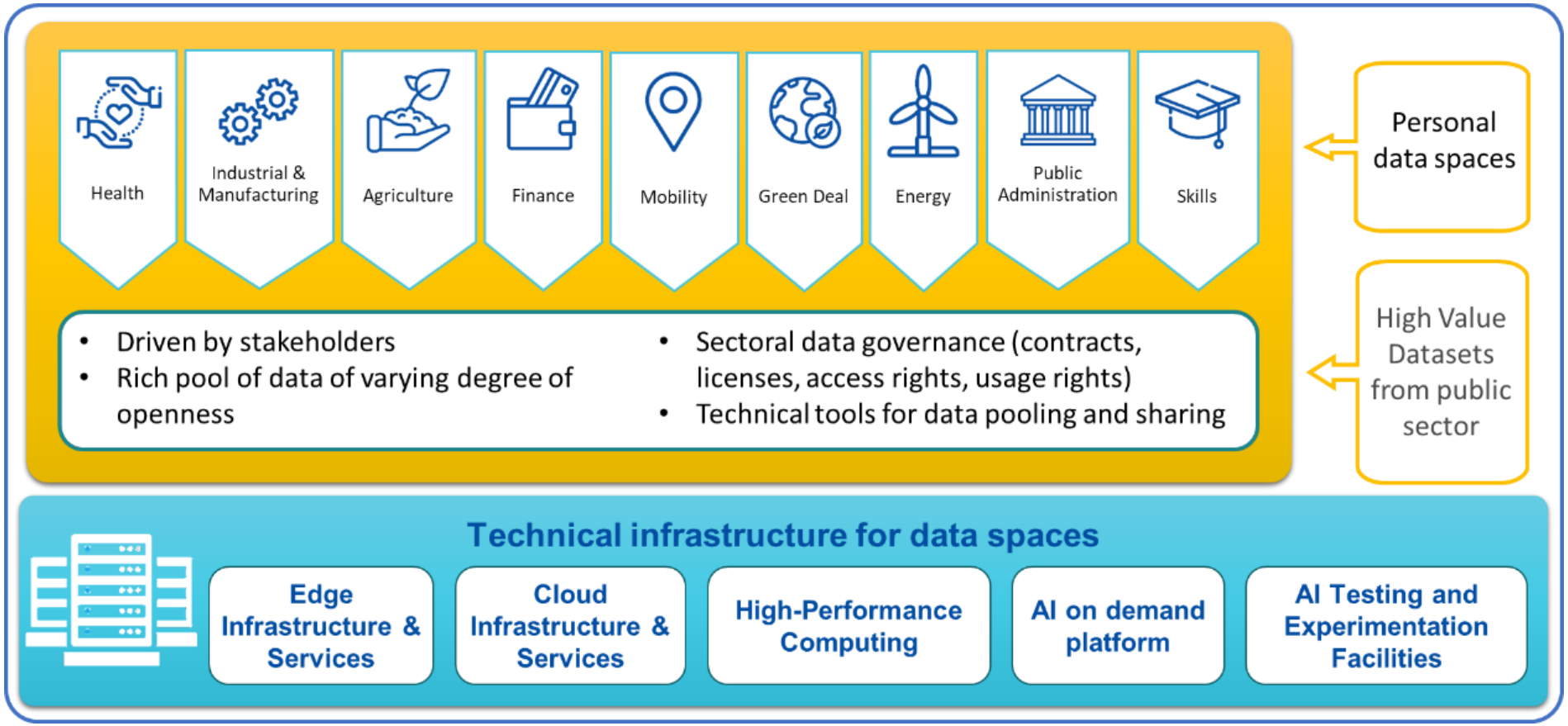


O1. study methods for improving a priori data and model searches, using metadata from data sets and models.
O2. create relevant indicators to define whether a set of information, data or models, is qualitative in a simulation.
O3. propose a data-model matching method using ontologies.

BEYOND OPEN DATA: DATA SPACES

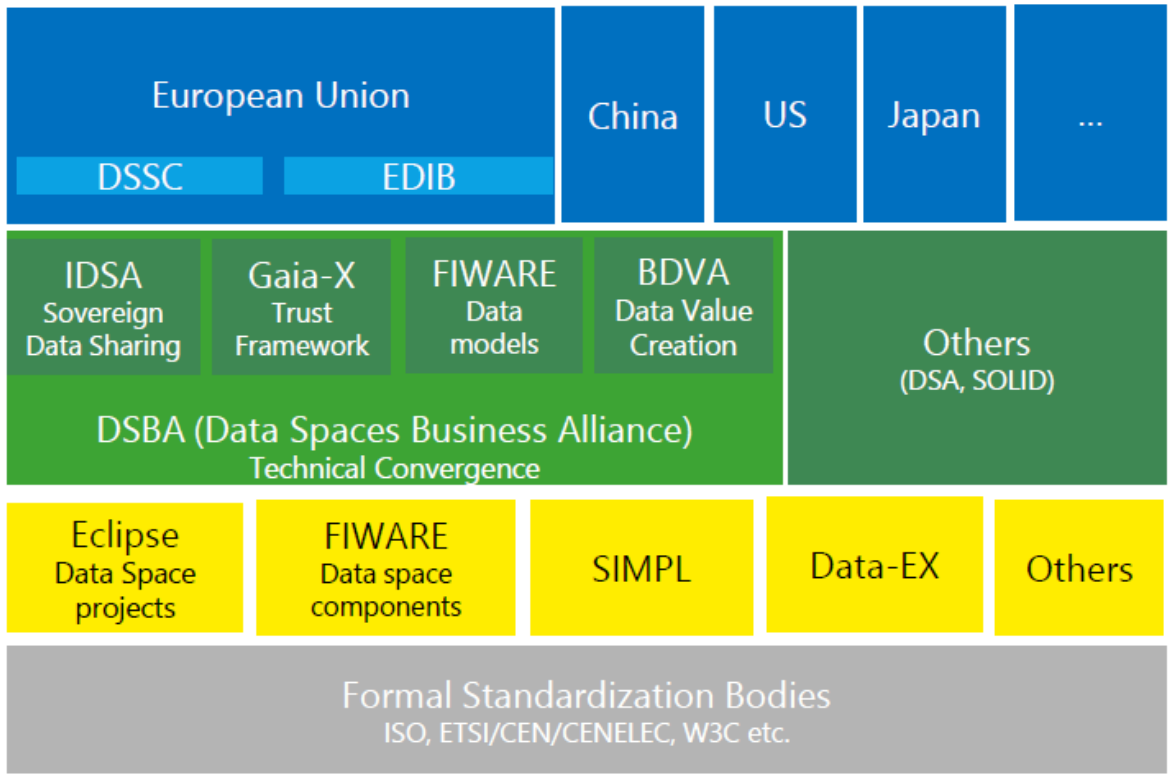
Common European data spaces

"purpose or sector specific or cross-sectoral interoperable frameworks for common standards and practices to share or jointly process data for, inter alia, the development of new products and services, scientific research or civil society initiatives." [source: Data Act]



https://digital-strategy.ec.europa.eu/en/library/building-data-economy-brochure

Regulatory, business, and technical foundation for Data Spaces within the Edge-Cloud-Continuum



Data regulations in economic regions

Data strategies implementation

User requirements, Voice of the communities, coordinate technical specs and business requirements, support to "business design"

Technical implementation driven by OSS, place for the developer communities

Long-term investment security, adoption support etc. through norms and standards

https://data-spaces-business-alliance.eu/download/33968/

Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

The Data Spaces Business Alliance

Unleashing the European Data Economy

https://data-spaces-business-alliance.eu/

Accelerating Business transformation in the Data Economy

The Data Spaces Business Alliance (DSBA) accelerates business transformation in the data economy. It's the first initiative of its kind, uniting industry players to realize a data-driven future in which organizations and individuals can unlock the full value of their data.

Data spaces are key to achieving sovereign, interoperable and trustworthy data-sharing across businesses and societies – a key step to the data economy of the future. The Alliance embraces this reality, converging the best skills, assets, and experience in Europe into a one-stop-shop for data spaces, from inception to deployment.

The Data Spaces Business Alliance are Gaia-X European Association for Data and Cloud AISBL, the Big Data Value Association (BDVA), FIWARE Foundation, and the International Data Spaces Association (IDSA). Together they represent 1,000+ leading key industry players, associations, research organizations, innovators, and policymakers worldwide. With this cross-industry expertise, resources and know-how, the Alliance drives awareness, evangelizes technology, shapes standards, and enables integration across industries.

The Data Spaces Business Alliance

Unleashing the European Data Economy

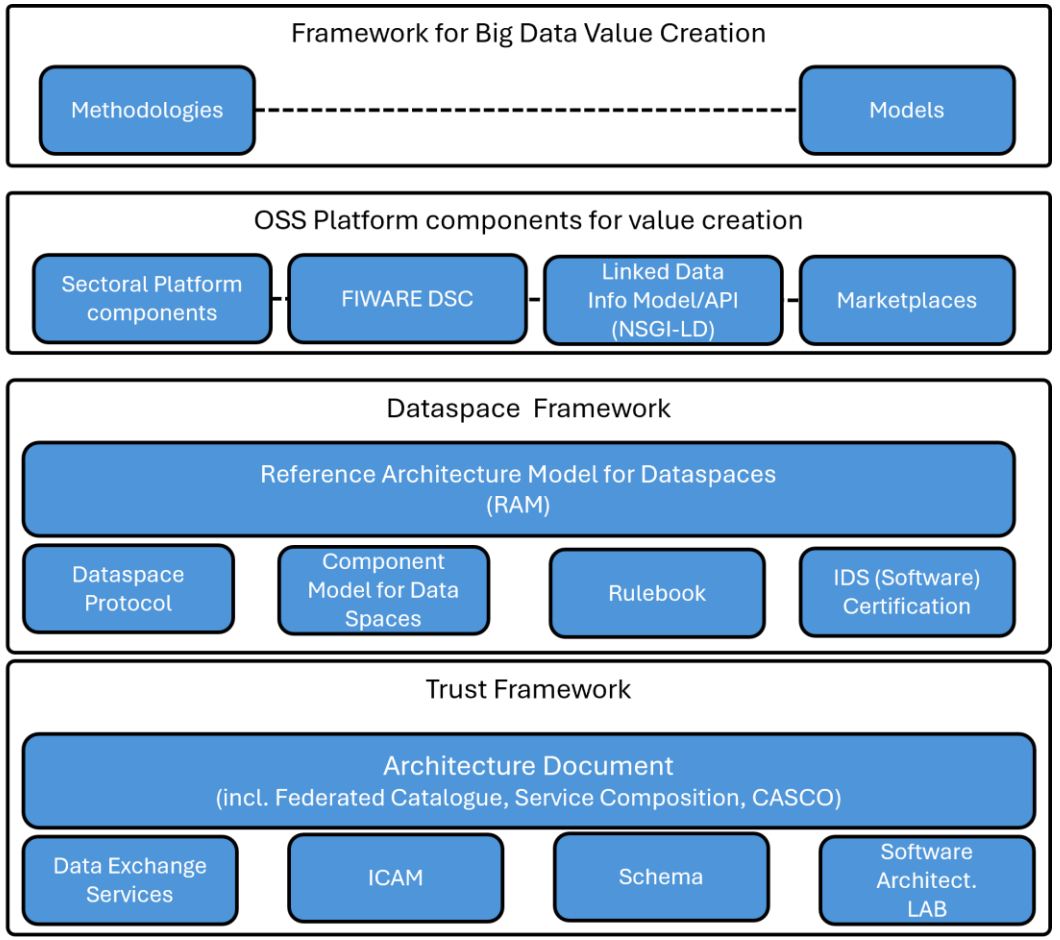
BDV BIG DATA VALUE ASSOCIATION

FIWARE Open APIs for Open Minds

INTERNATIONAL DATA SPACES ASSOCIATION

gaia-x

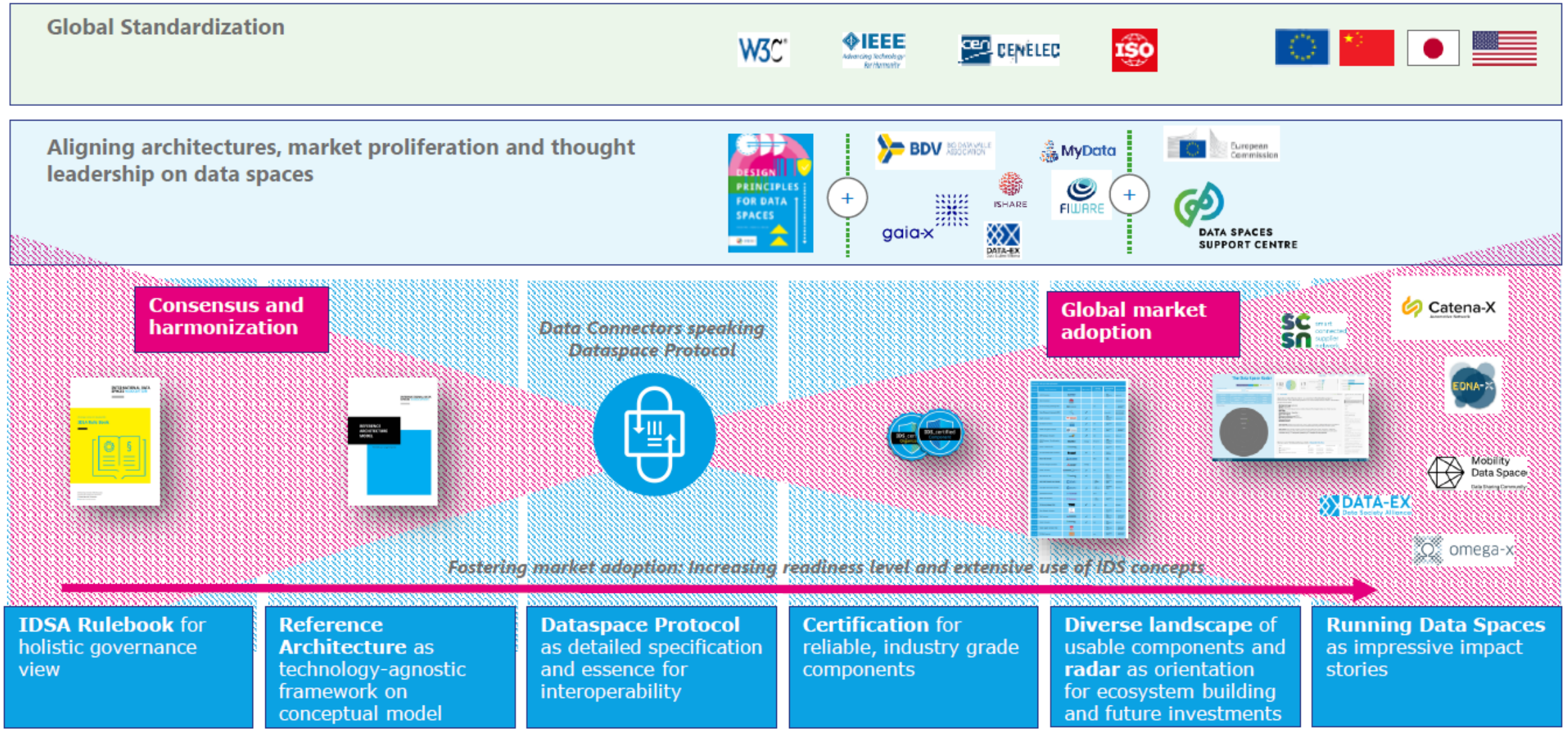
https://data-spaces-business-alliance.eu/



A holistic approach to bring data spaces to global scale

INTERNATIONAL DATA SPACES ASSOCIATION

IDSA on its way to a global standard – with the dataspace protocol in its core



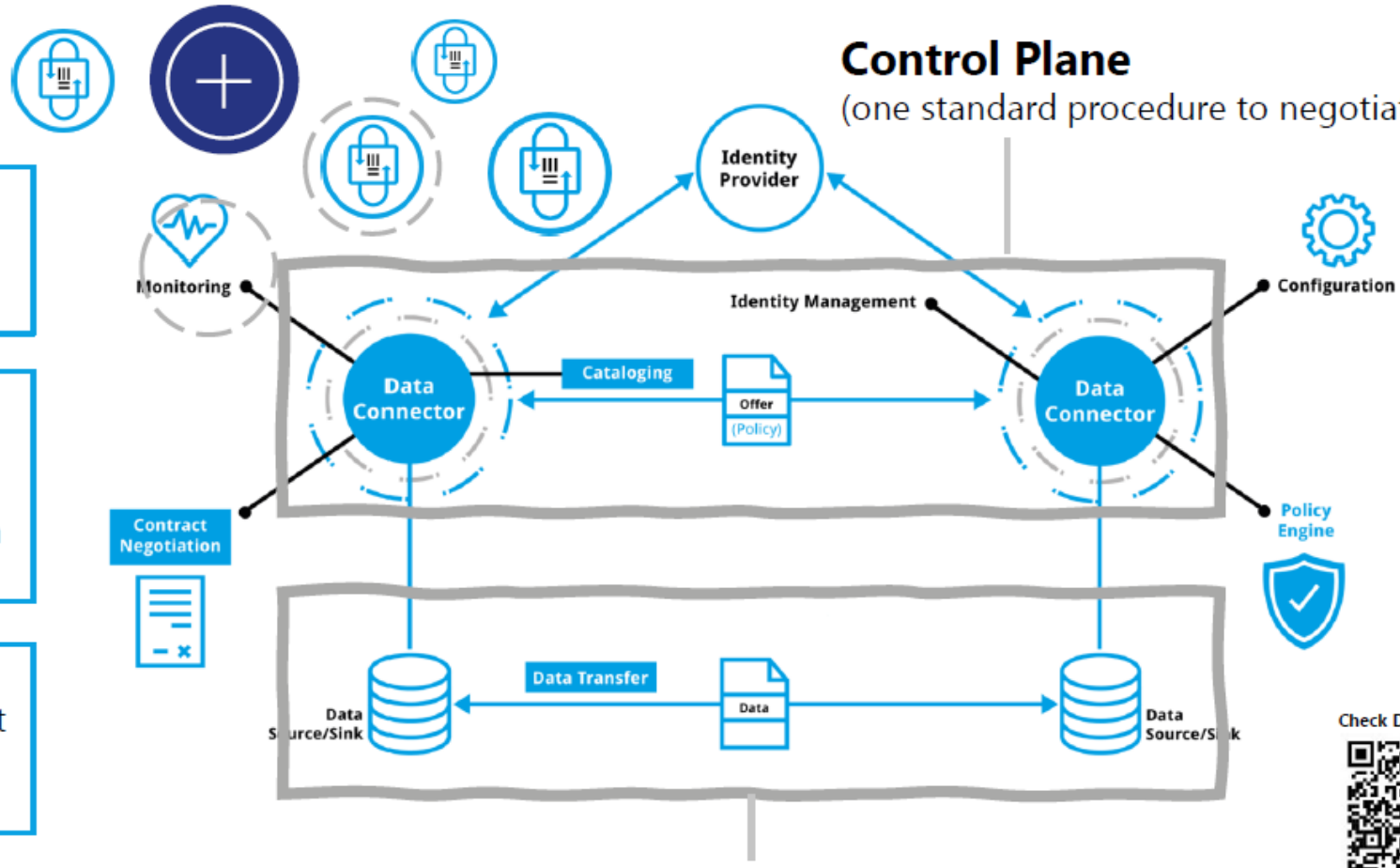
Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

Dataspace Protocol V1.0 → ISO Standard

Enables standardized data exchange across different data space instances.

INTERNATIONAL DATA SPACES ASSOCIATION

- Ensures standardized data exchange mechanism between different frameworks, products, or services.
- Rrovides the needed schemas and protocols for cataloging data, negotiating contracts and usage agreements, and accessing data within a data space.
- Organizations using this protocol can align with industry standards, foster best practices, and unlock new data-driven business models and opportunities.



Check Dataspace Protocol:

16

Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

Dataspace Protocol V1.0 → ISO Standard

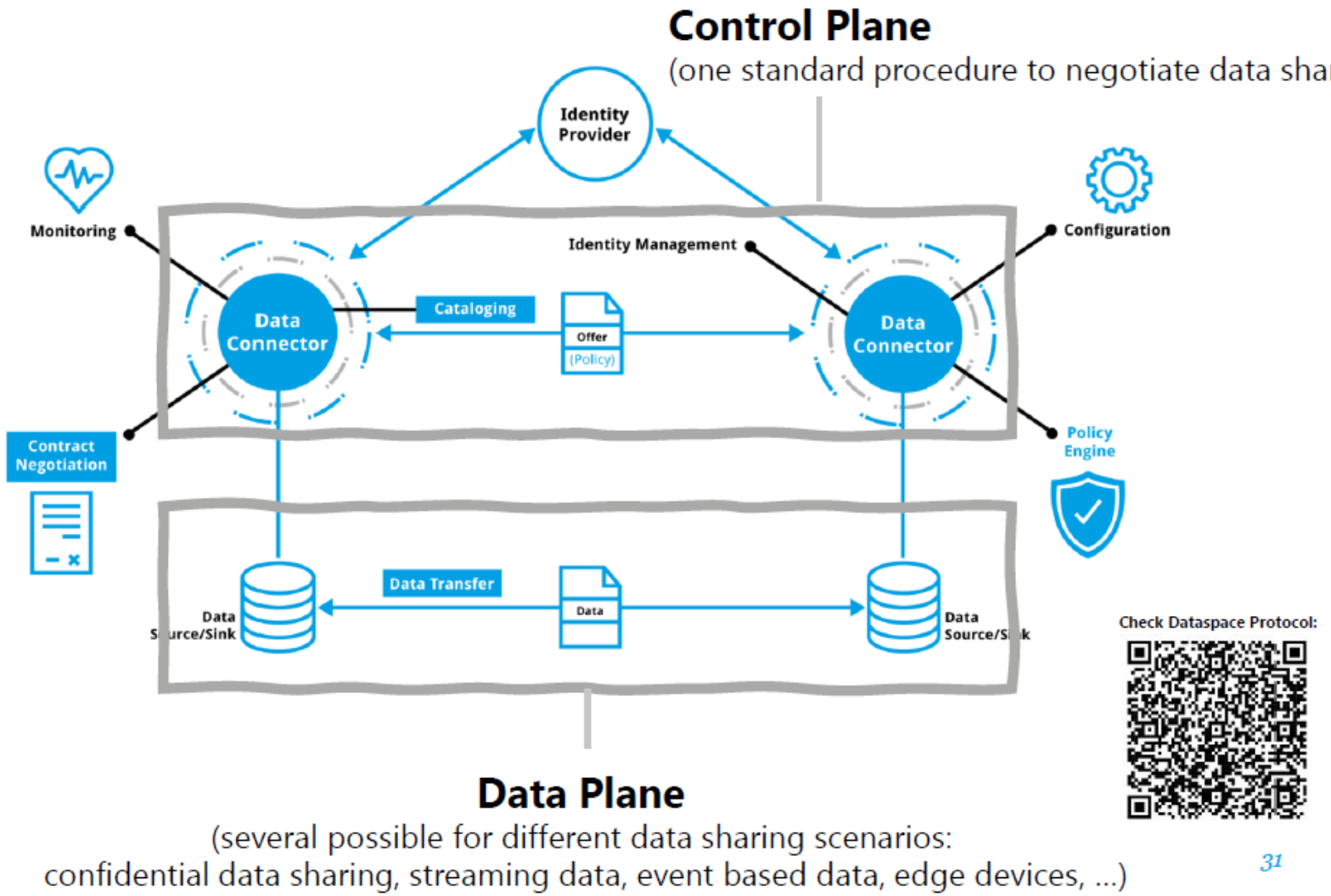
Enables standardized data exchange across different data space instances.

INTERNATIONAL DATA SPACES ASSOCIATION

Control Plane decides who can access the data and how.

Data Plane is where the action (data sharing) happens.

Conceptually divided, can be combined practically



Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

Make the connection and enable data economy

The key to data spaces is the data connector

INTERNATIONAL DATA SPACES ASSOCIATION

- **Connects participants in a data space** – to share, utilize, benefit from data.
- Ensures **trust through IDS Certification** and **cyber security** assessment.
- Connects to **trust frameworks** and **identity management**
- Includes **identity & policy management**, ensures **data usage control**.
- Guarantees **interoperability**.
- Understands and enforces **data usage policies**.
- **Master** for other connectors of diverse feature sets.

Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

Standardized Data Exchange

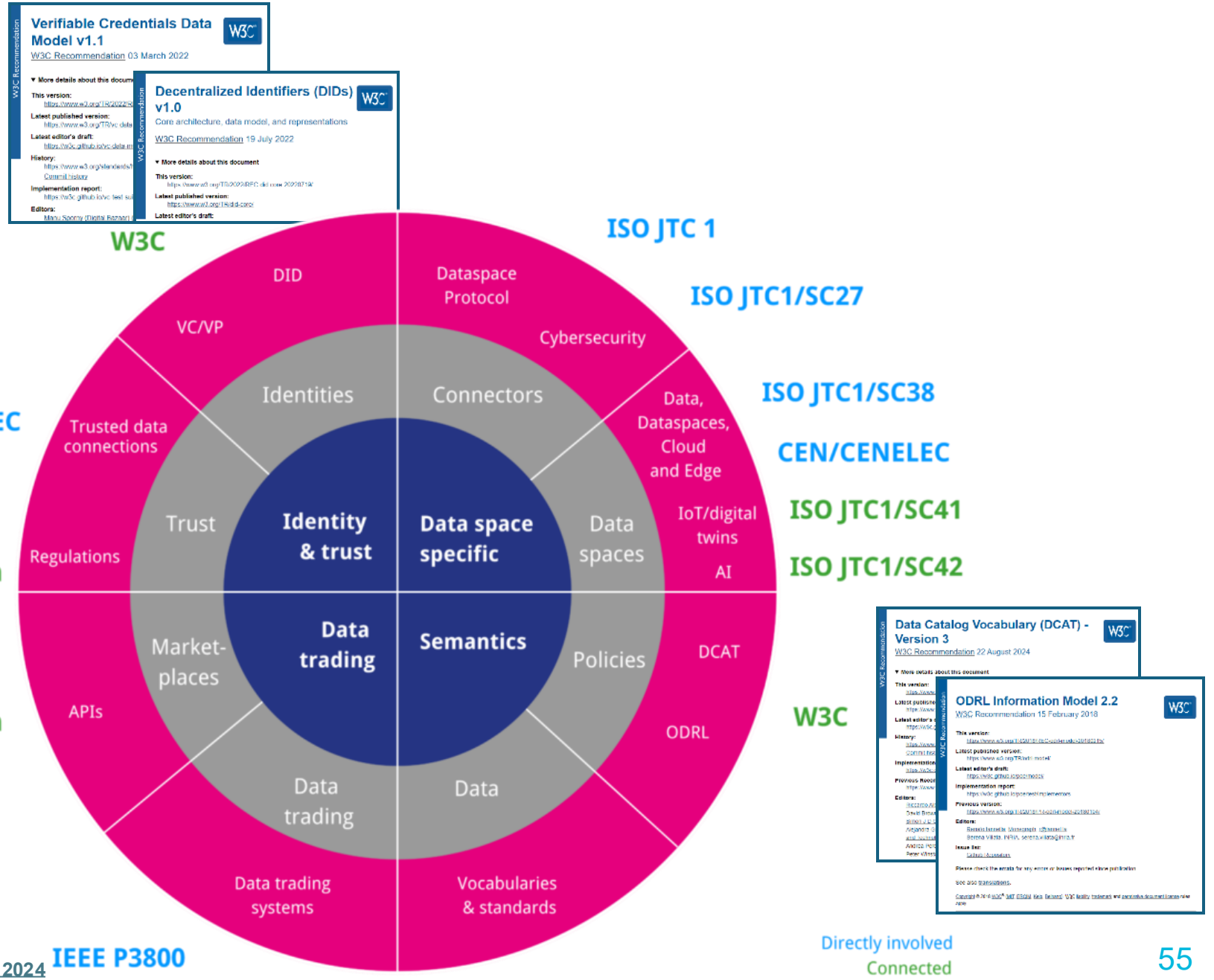
What does this mean? How does Dataspace Protocol ensure that?

INTERNATIONAL DATA SPACES ASSOCIATION

	Catalog	Contract Negotiation	Transfer Process
What happens?	A public transportation authority decides to share its transit schedules with app developers.	An app developer wants to use these schedules to create a route planning application.	Once the agreement is in place, the actual data transfer begins.
Problem	Inconsistent data formats for schedules across different platforms.	Need for clear terms regarding the use and distribution of the transit data.	Ensuring secure, efficient, and reliable transfer of transit data.
Role of DSP	Standardizes the format for publishing transit schedules.	Facilitates agreement on data usage terms and conditions.	Manages the secure and efficient transfer of the agreed-upon data.
Specification Example	Data provider publishes schedules using *'DCAT Catalogs'* and sets access rules with *'ODRL Policies'*.	Developer and authority negotiate using *'Contract Offer'* messages, leading to a *'Contract Agreement'*.	Data transfer is executed through *'Connector-to-Connector Communication'* and *'Data Transfer Requests'*.

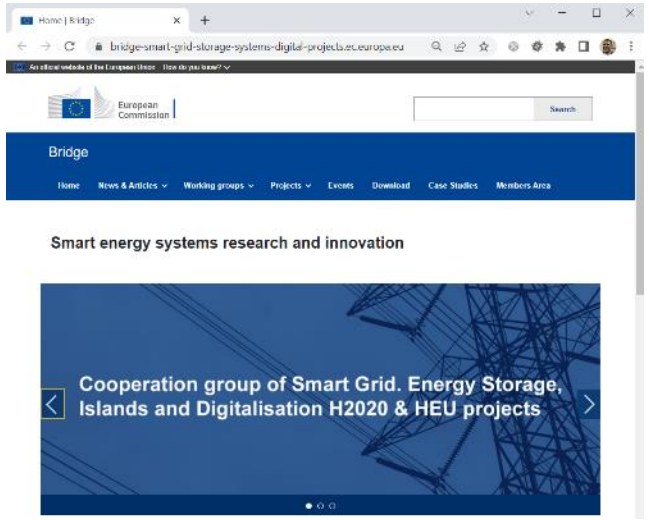
Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

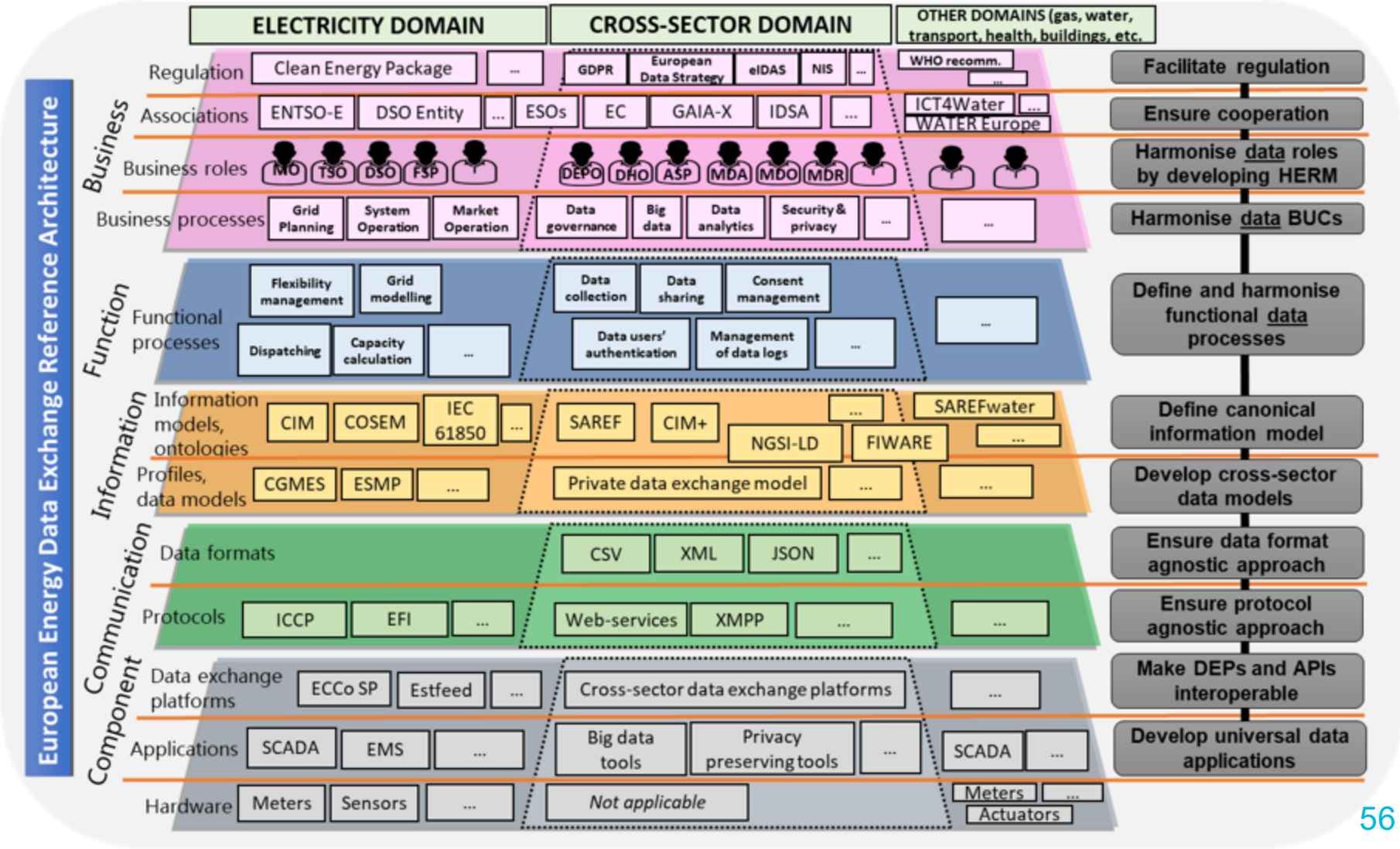
Standardization activities



Adapted from a presentation by L. Nagel at the Data Spaces Symposium 2024

EU Bridge – Use Cases and Reference architecture for data exchange in the *EU Energy Data Space*





INTERNATIONAL DATA SPACES ASSOCIATION

Data Spaces Radar

The Data Spaces Radar
Map View

Radar View | Chart View | **Map View** | Table View | Building Blocks

Geographical Focus
- Europe
- European
- International
- Local
- National (scope is onl...

Countries
- Austria
- Belgium
- Belgium, Finland
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- Finland, Spain
- France

Filter results:

All EU Member States

Community of Practice | Data Space | Use Case

Agricult...	Agricult...	Autom...	Autom...
Buildin...	Built E...	Cross ...	Cultura...
Energy	Finance	Geoinf...	Green ...
Green ...	Health	Langua...	Logistics
Manufa...	Media	Mobility	Open s...

Data Space: Tout

Development Stage: Tout

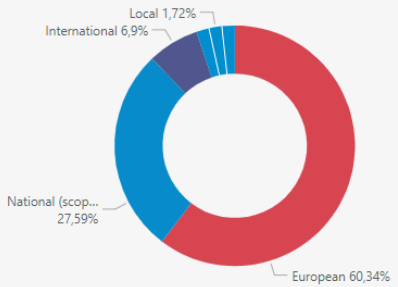
Source of Funding: Tout

Reference Architecture Used: Tout

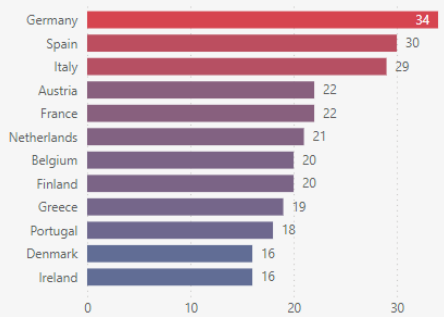
Period: 01/01/2018 – 07/01/2024

Freeform Search: Search

Number of Entries
par Geographical Focus



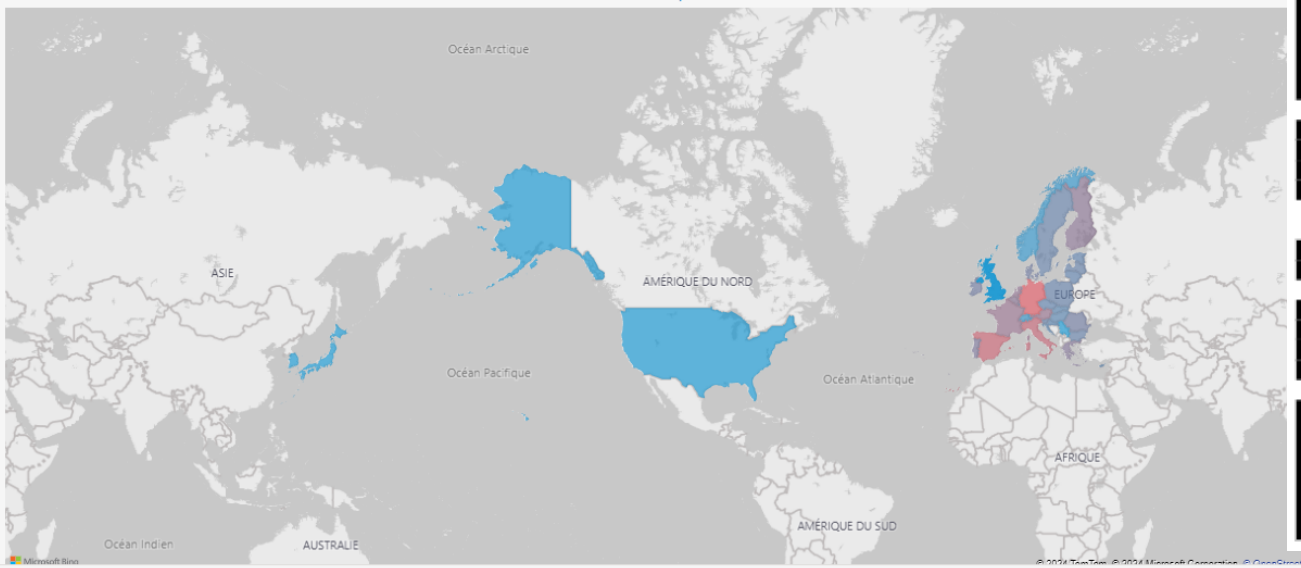
Number of Entries
par Countries



Relevant Entries

Name	Type
Würth C-Part Supply	Use Case
Wind Energy Generation Data Space	Use Case
Wind and Solar Assets modeling	Use Case
WHIRLPOOL After -Sale Consumer Services powered by Predictive Analytics	Use Case
Volkswagen Autoeuropa Intra-logistics process planning powered by simulation	Use Case
VELES Project	Use Case

Map View



https://www.dataspaces-radar.org/

Example of a project on the Energy Data Space

Omega-X

Common Semantic Data Model & Semantic Interoperability

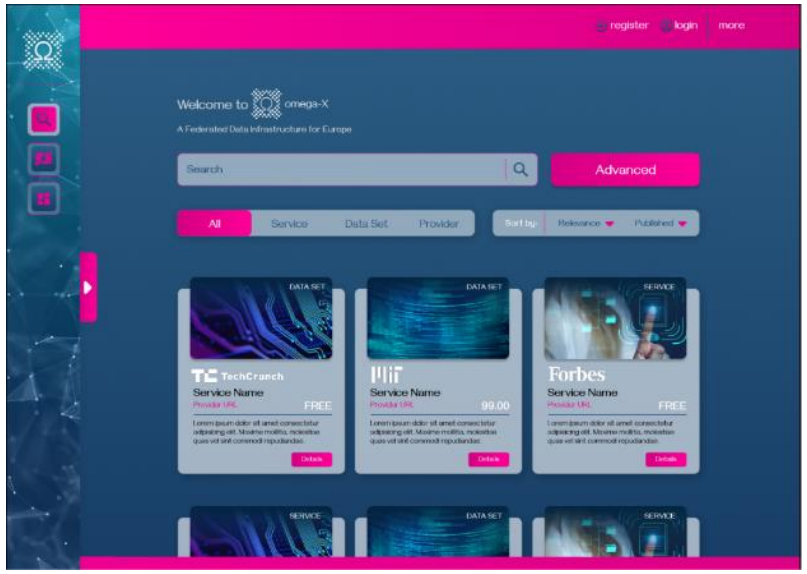
EDF

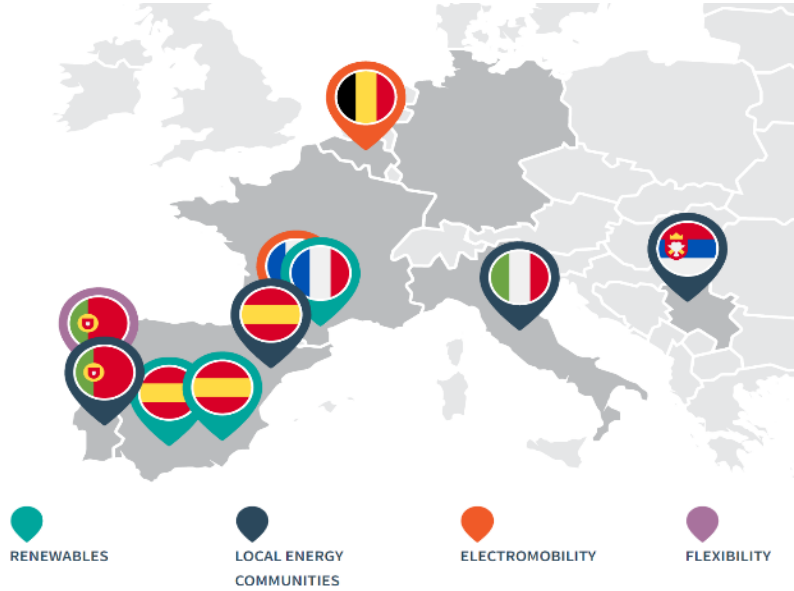
Orchestrating an interoperable sovereign federated Multi-vector Energy data space built on open standards and ready for GAia-X

Funded by the European Union

omega-x

Orchestrating an interoperable sovereign federated Multi-vector Energy Data Space built on open standards and ready for GAia-X





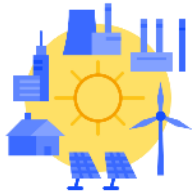
Renewables

- PV Cleaning Advisor
- Fraud detection
- Tracking algorithm check

Local Energy Com

- Electrical Losses Detection
- Gamification for electrical energy savings
- Water losses detection

Flexibility



- Passive consumption baseline prediction service
- Grid observability and network analysis

Electromobility

- Roaming of booking services
- Roaming of solar energy
- Day-ahead availability forecast

Funded by the European Union

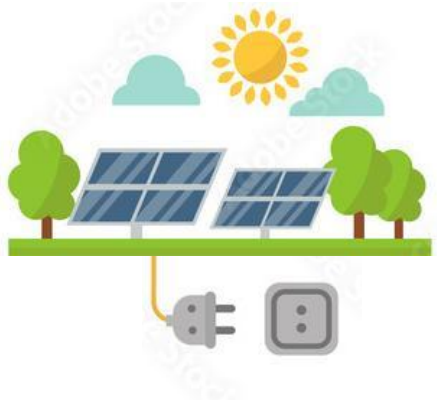
Context

A key challenge for data spaces is overcoming **heterogeneity**. Datasets are provided by multiple entities, in variable: formats, languages, data models and structures.

Multi-factor heterogeneity leads to difficulties in data **discoverability** and **understanding**, challenging an optimal use of data, for expected use cases.

Case study

Renewable Generator Base Training: this case study aims to make predictions about active energy production, considering meteorogical indicators (temperature, irradiance). To this aim, machine learning algorithms require processing relevant historical data.



Dataset that includes electrical photovoltaic production data along with meteorological indicators, in Portugal, from the year 2021

Funded by the European Union

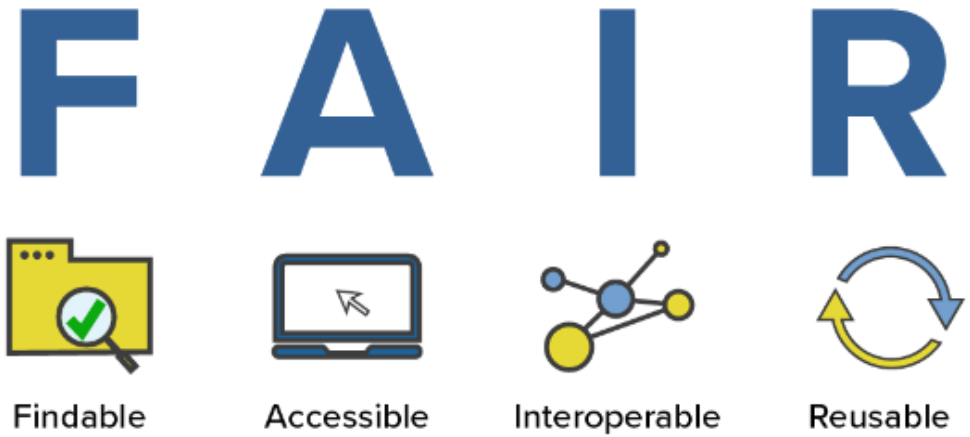
Scope

To ensure that data providers and service providers in Omega-X data space have a common understanding of shared datasets.

Drivers

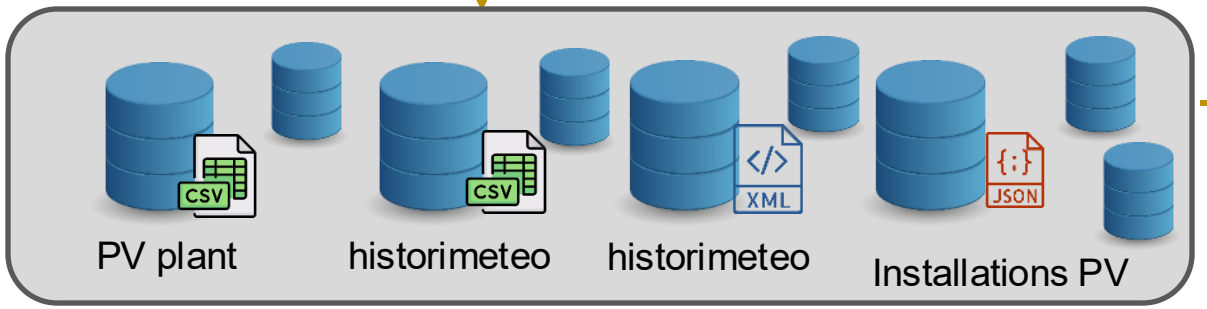
✓ Conformance to energy domain standards (IEC CIM, IEC 61850,…)

✓Support of FAIR data principles (Energy Data Act)



Context

Electrical photovoltaic production , meteorological indicators, Portugal, 2021



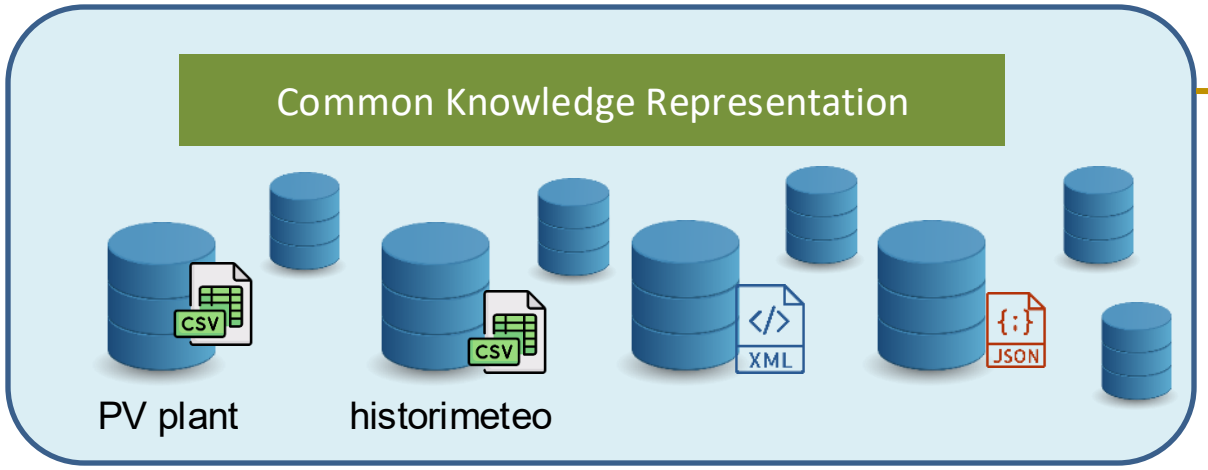
Returned Datasets

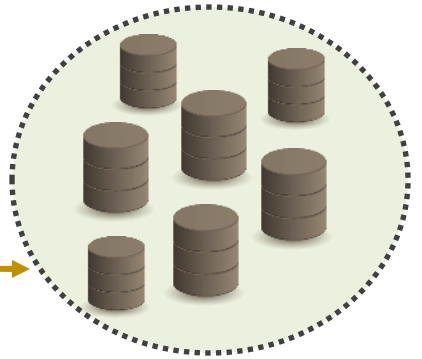
- Does this set exhaustively cover search query needs ?
- For each dataset, to which extent I can trust its quality?
- To what correspond each attribute in the dataset ?
- What encoding is used for energy ?
- Do I need to align units for active energy ?

Funded by the European Union

Context

Electrical photovoltaic production , meteorological indicators, Portugal, 2021





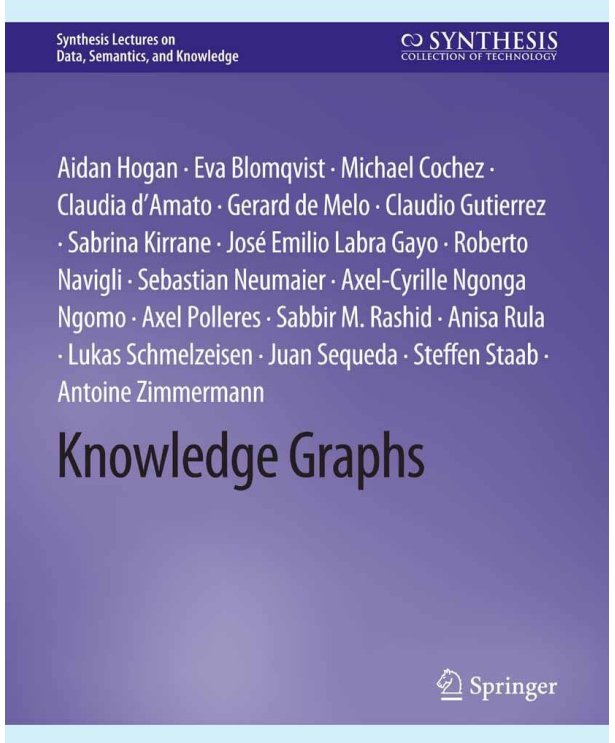
Returned Datasets

- Wider **Discoverability**
- A common **data understanding**
- **Query** each data set to apply filters
- Derive **implicit** knowledge
- Domain knowledge-based **Validation**

Funded by the European Union

Vocabularies, Taxonomies, Ontologies, … ?

Data Graphs



Section based on A. Hogan et. al., 2020. Knowledge graphs. arXiv:2003.02320 (version html / https://kgbook.org) and on Springer

Why graphs ?

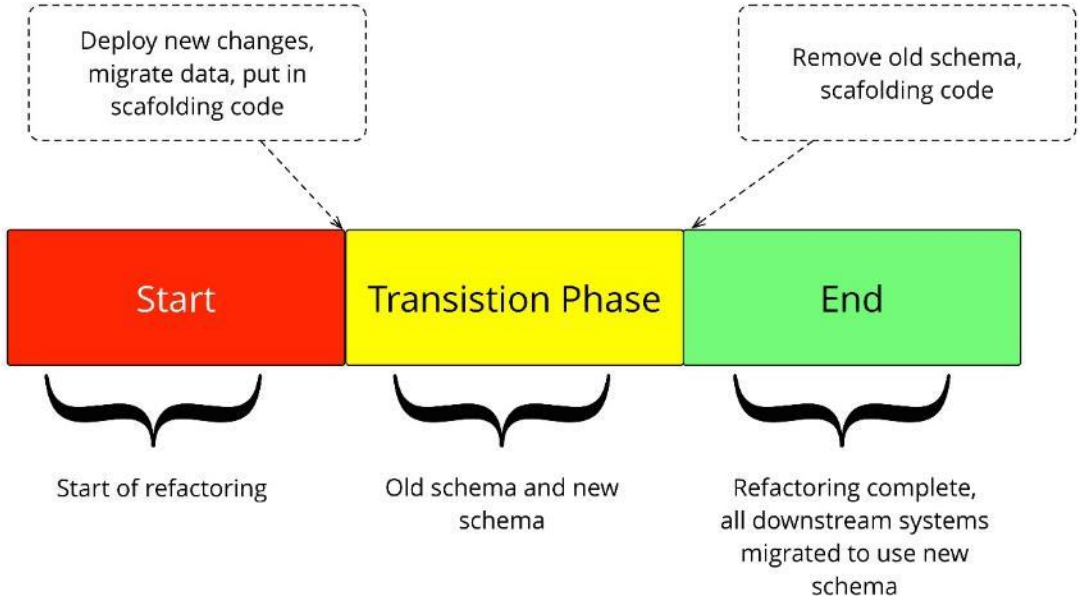
Evolution of a relational database schema ?

Event(<u>name</u>, venue, type, <u>start</u>, end)

but...
« an event can have different names in different languages »
« an event can have many venues »
« an even can have more than one types »
...

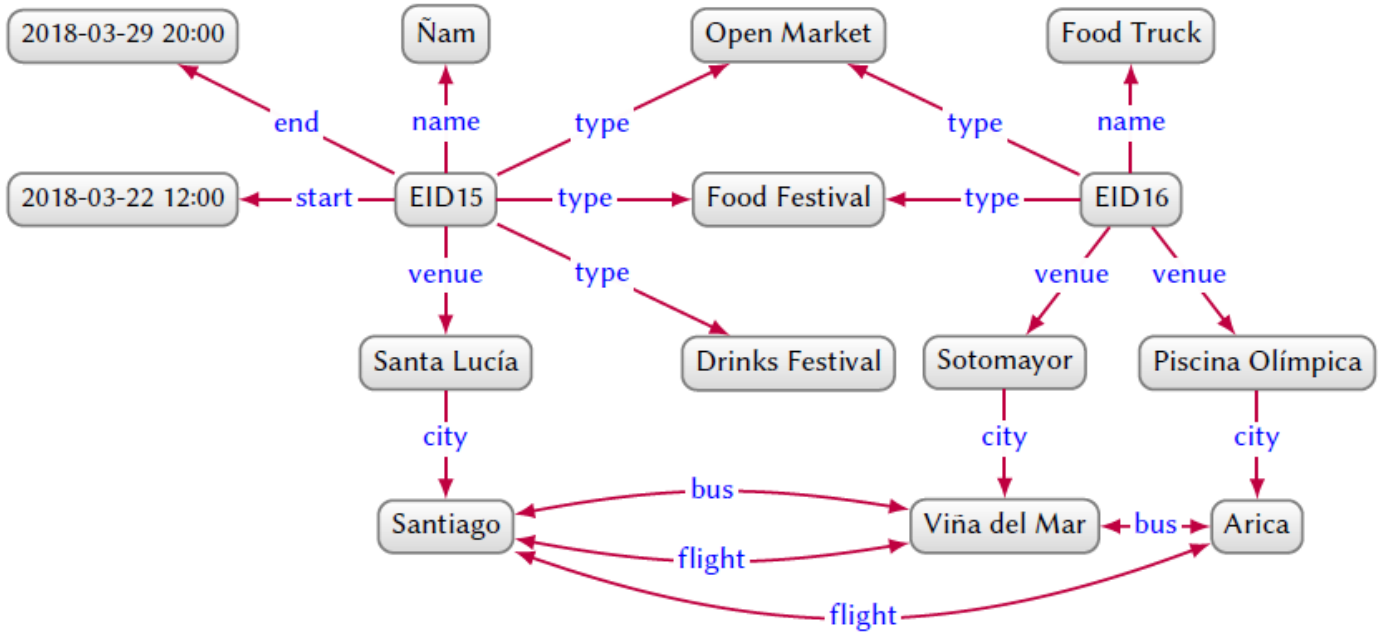
EventName(<u>id</u>, <u>name</u>), EventStart(<u>id</u>, start), EventEnd(<u>id</u>, end),
EventVenue(<u>id</u>, <u>venue</u>), EventType(<u>id</u>, type)

Binary relations between entities = graph



source: Evolutionary Database Design
https://martinfowler.com/articles/evodb.html

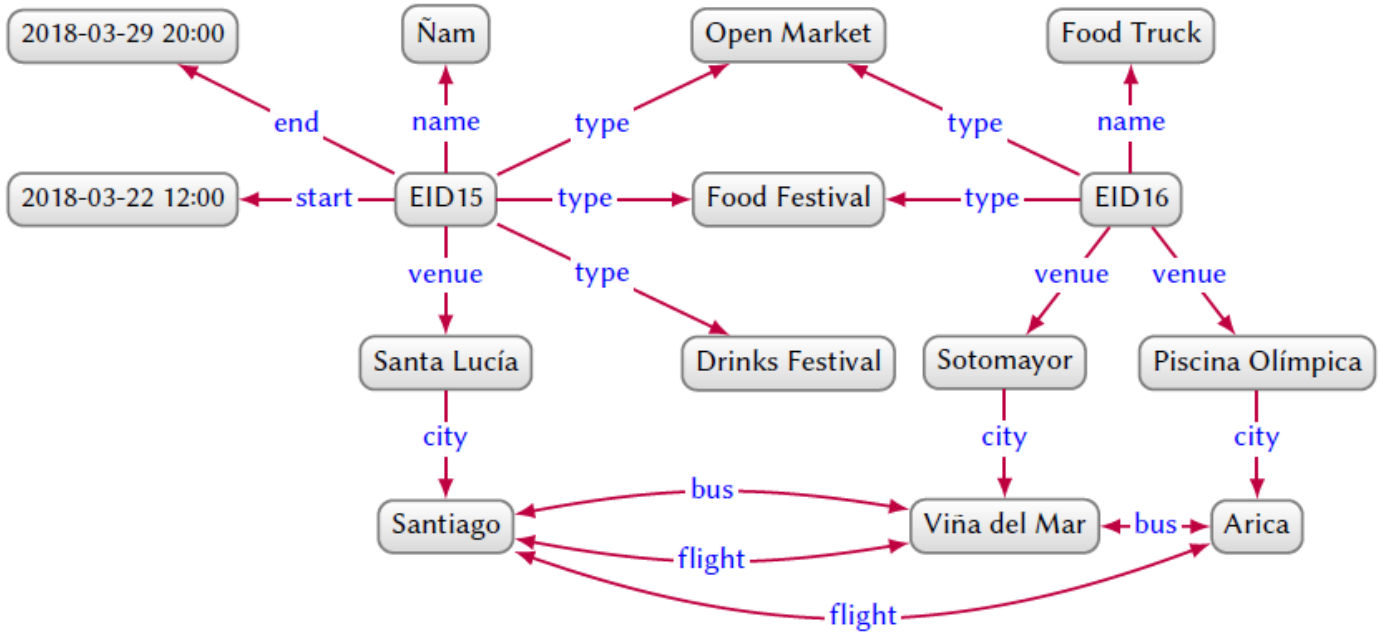
Directed edge-labelled graph



Definition B.1 (Directed edge-labelled graph). A *directed edge-labelled graph* is a tuple $G := (V, E, L)$, where $V \subseteq \mathbf{Con}$ is a set of nodes, $L \subseteq \mathbf{Con}$ is a set of edge labels, and $E \subseteq V \times L \times V$ is a set of edges.

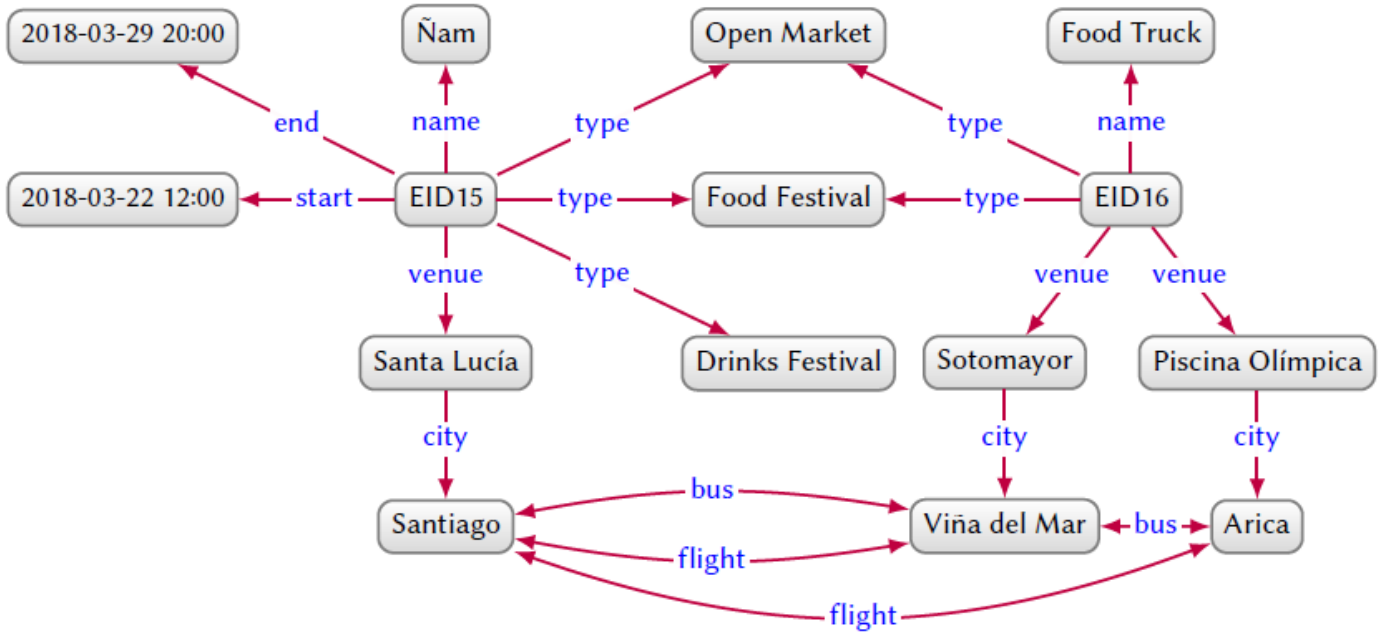
Example B.2. In reference to Figure 1, the set of nodes V has 15 elements, including `Arica`, `EID16`, etc. The set of edges E has 23 triples, including (`Arica`,`flight`,`Santiago`). Bidirectional edges are represented with two edges. The set of edge labels L has 8 elements, including `start`, `flight`, etc.

Directed edge-labelled graph



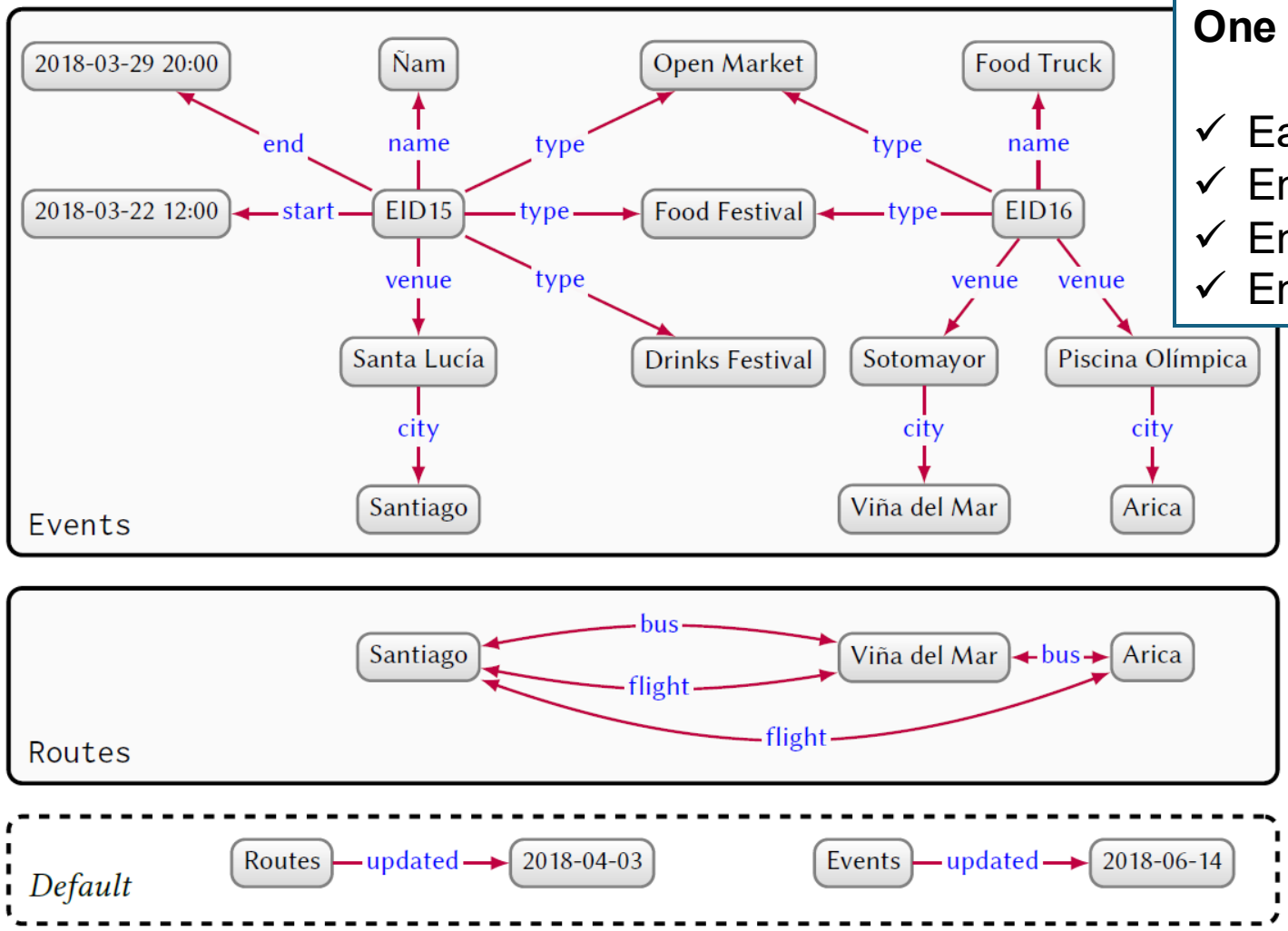
nouvelle information: ajout d'entités et d'arcs
absence d'information: pas d'arc
facile à faire évoluer
facile à intégrer de nouvelles données
pas de hiérarchisation (opp. XML, JSON)

Directed edge-labelled graph



For example W3C Resource Description Framework (RDF)
- subject is an *Internationalised Resource Identifier* (IRI) *or a* blank node
- edge is an IRI
- object is an IRI or blank node or literal (character string + datatype IRI)
- a literal has not outgoing edges

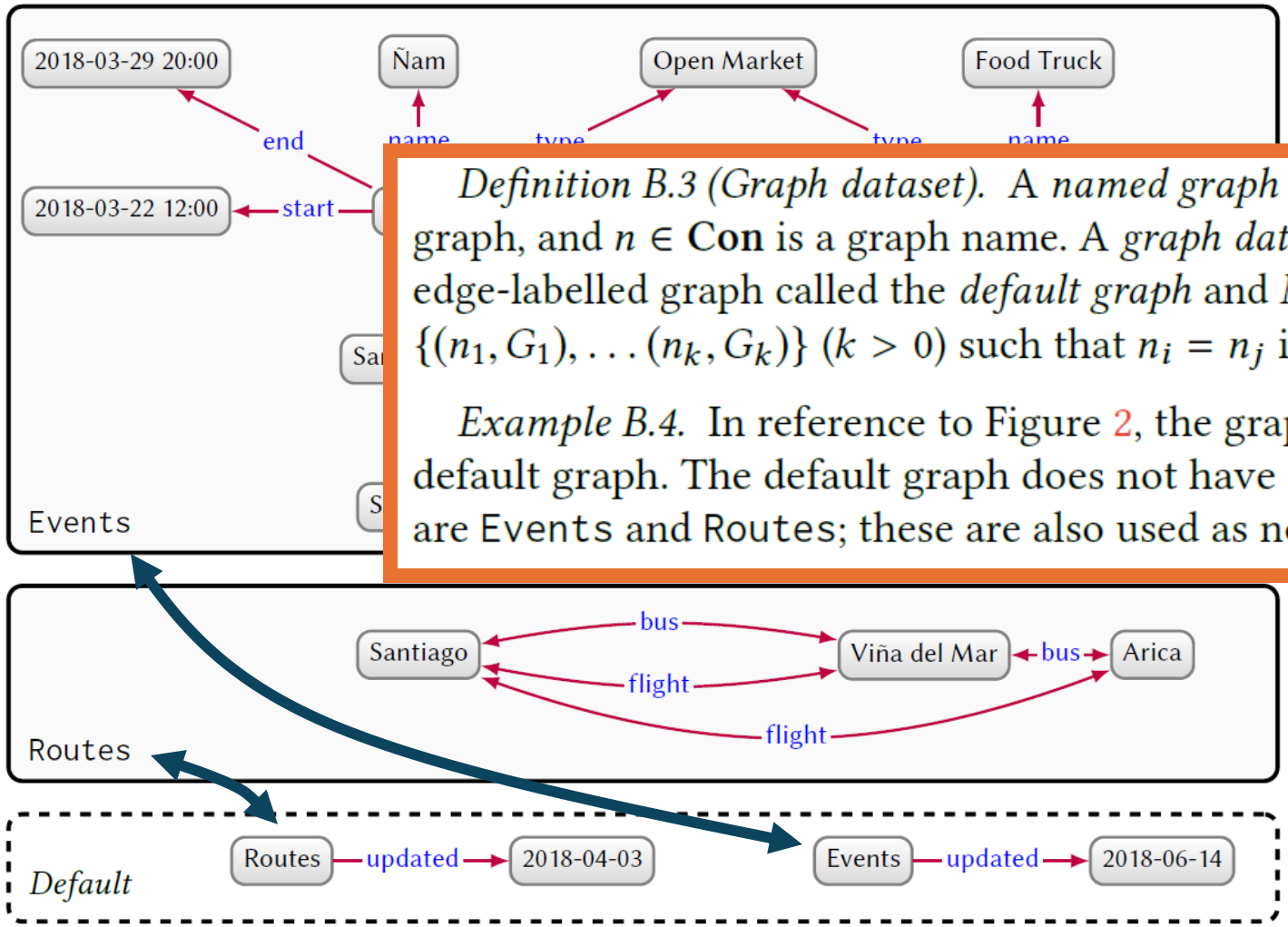
Variant: Graph dataset



One default graph and named graphs

- ✓ Easier to manage than one single monolithic graph
- ✓ Enables to separate data by source
- ✓ Enables to model context
- ✓ Enables to describe differently the same entities

Variant: Graph dataset

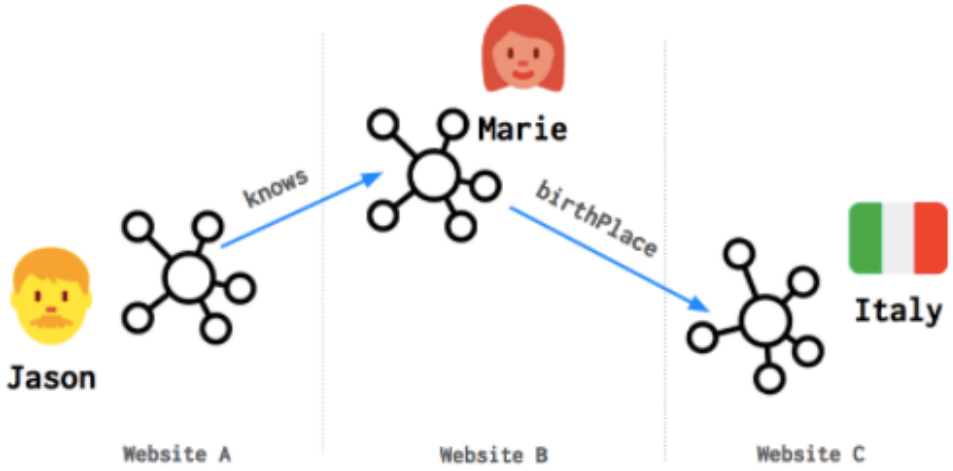


Definition B.3 (Graph dataset). A *named graph* is a pair (n, G) where G is a directed edge-labelled graph, and $n \in \mathbf{Con}$ is a graph name. A *graph dataset* is a pair $D := (G_D, N)$ where G_D is a directed edge-labelled graph called the *default graph* and N is either the empty set or a set of named graphs $\{(n_1, G_1), \ldots (n_k, G_k)\}$ $(k > 0)$ such that $n_i = n_j$ if and only if $i = j$ $(1 \leq i \leq k, 1 \leq j \leq k)$.

Example B.4. In reference to Figure 2, the graph dataset D consists of two named graphs and a default graph. The default graph does not have a name associated with it. The two graph names are `Events` and `Routes`; these are also used as nodes in the default graph.

Variant: Graph dataset

Example: the Web of Linked Data



source: article on Linked Data by WordLift https://wordlift.io/blog/en/entity/linked-data/

Support: *RDF triplestores*

Ontotext GraphDB

Ontotext GraphDB is a graph database and knowledge discovery tool compliant with RDF and SPARQL and available as a high-availability cluster. Ontotext GraphDB is used in various European research projects. As of April 2021, Graph DB is ranked as the 4th most -popular RDF store and 6th most-popular Graph DBMS system. Wikipedia

Stardog

Stardog is a **commercial RDF database**: insanely fast SPARQL query, transactions, and world-class OWL reasoning support. Retrieved from "https://www.w3.org/2001/sw/wiki/index.php?title=Stardog&oldid=4920" Mar 27, 2015

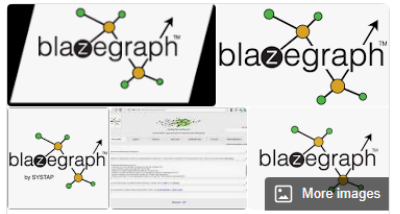
https://www.w3.org › wiki › Stardog
Stardog - Semantic Web Standards

AllegroGraph

AllegroGraph is a closed source triplestore which is designed to store RDF triples, a standard format for Linked Data. It also operates as a document store designed for storing, retrieving and managing document-oriented information, in JSON-LD format. Wikipedia

Developer: Franz

License: Proprietary commercial software



Blazegraph

Blazegraph is a triplestore and graph database, which is used in the Wikidata SPARQL endpoint. Wikipedia

License: GNU GPL (version 2)

Preview release: 2.1.6rc / 3 February 2020

Stable release: 2.1.5 / 19 March 2019

Alternative: Property Graphs

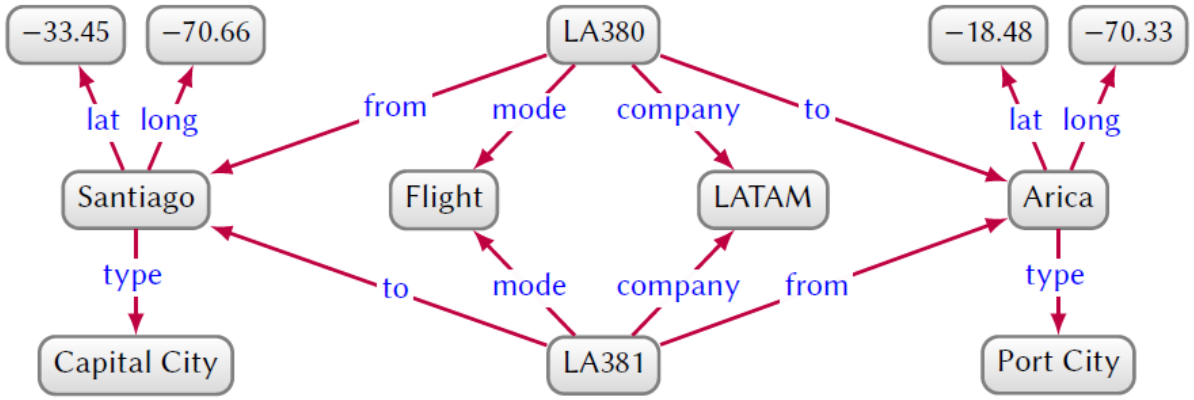


Fig. 3. Directed edge-labelled graph with companies offering flights between Santiago and Arica

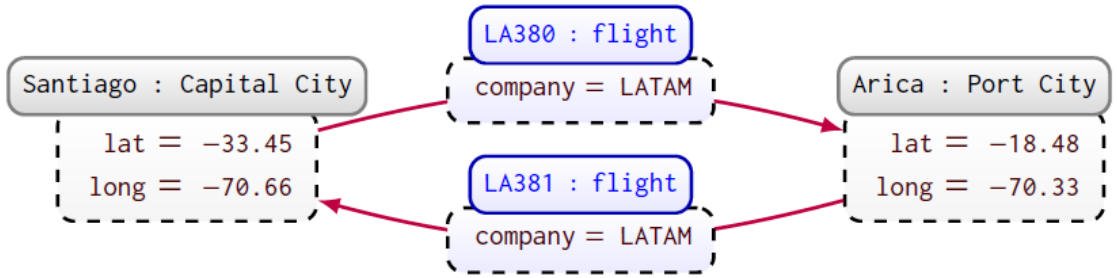


Fig. 4. Property graph with companies offering flights between Santiago and Arica

Each entity or relationship is labelled **et** has a set of key-value pairs

✓ More flexible
✓ Popular in commercial graph databases

Alternative: Property Graphs

Definition B.5 (Property graph). A *property graph* is a tuple $G := (V, E, L, P, U, e, l, p)$, where $V \subseteq \mathbf{Con}$ is a set of node ids, $E \subseteq \mathbf{Con}$ is a set of edge ids, $L \subseteq \mathbf{Con}$ is a set of labels, $P \subseteq \mathbf{Con}$ is a set of properties, $U \subseteq \mathbf{Con}$ is a set of values, $e : E \to V \times V$ maps an edge id to a pair of node ids, $l : V \cup E \to 2^L$ maps a node or edge id to a set of labels, and $p : V \cup E \to 2^{P \times U}$ maps a node or edge id to a set of property–value pairs.

Example B.6. Returning to Figure 4:

- the set V contains `Santiago` and `Arica`;
- the set E contains `LA380` and `LA381`;
- the set L contains `Capital City`, `Port City`, and `flight`;
- the set P contains `lat`, `long`, and `company`;
- the set U contains -33.45, -70.66, `LATAM`, -18.48, and -70.33;
- the mapping e gives, for example, $e(\texttt{LA380}) = (\texttt{Santiago}, \texttt{Arica})$;
- the mapping l gives, for example, $l(\texttt{LA380}) = \{\texttt{flight}\}$ and $l(\texttt{Santiago}) = \{\texttt{Capital City}\}$;
- the mapping p gives, for example, $p(\texttt{Santiago}) = \{(\texttt{lat}, -33.45), (\texttt{long}, -70.66)\}$ and $p(\texttt{LA380}) = \{(\texttt{company}, \texttt{LATAM})\}$.

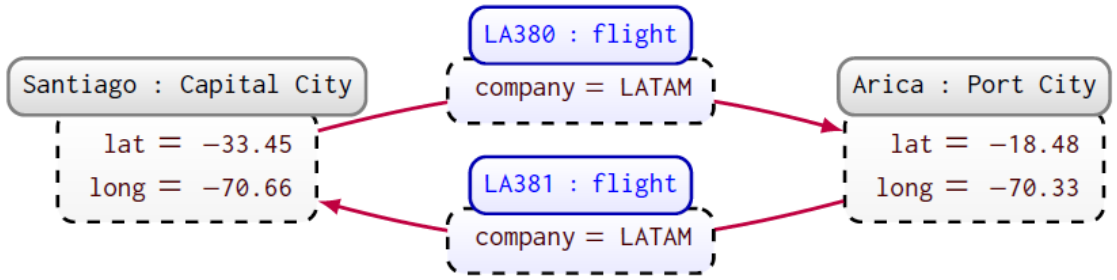
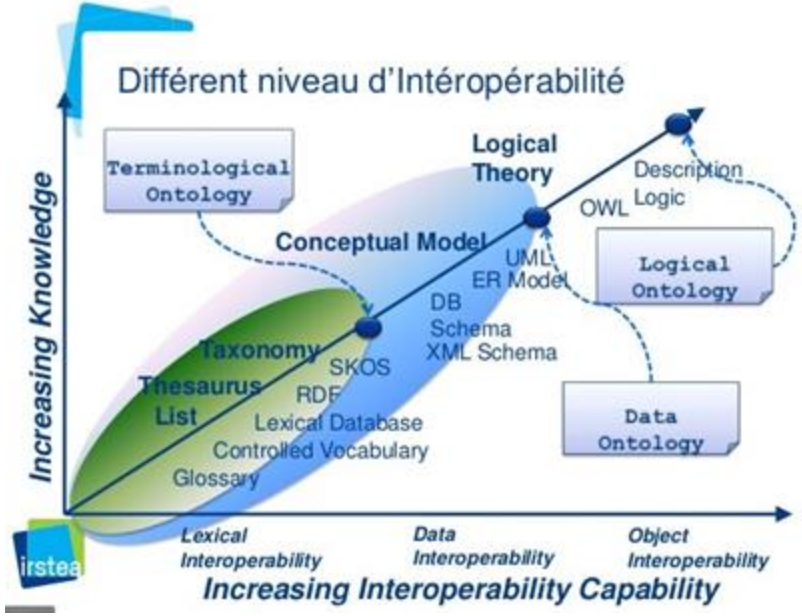


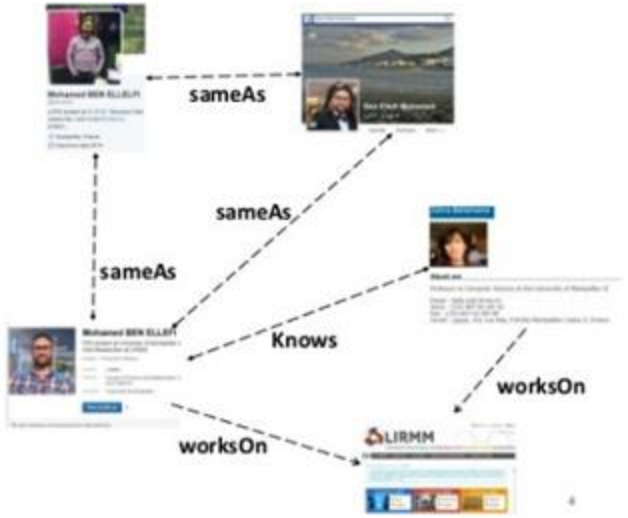
Fig. 4. Property graph with companies offering flights between Santiago and Arica

Data graph → knowledge graph

Schema, Identity, Context



source: https://fr.slideshare.net/croussey/skos-transformation



https://www.slideshare.net/MohamedBENELLEFI/profilebased-dataset-recommendation-for-rdf-data-linking

Data Graph -> Knowledge Graph

Knowledge graph = Data graph whose purpose is to accumulate and transmit real-world knowledge, whose nodes represent entities of interest, and whose arcs represent relationships between these entities.

Knowledge graph = Data graph enriched with representations of a schema, identity, context, ontology, validation rules, inference rules, etc.

Schema 1/3: *semantic* type

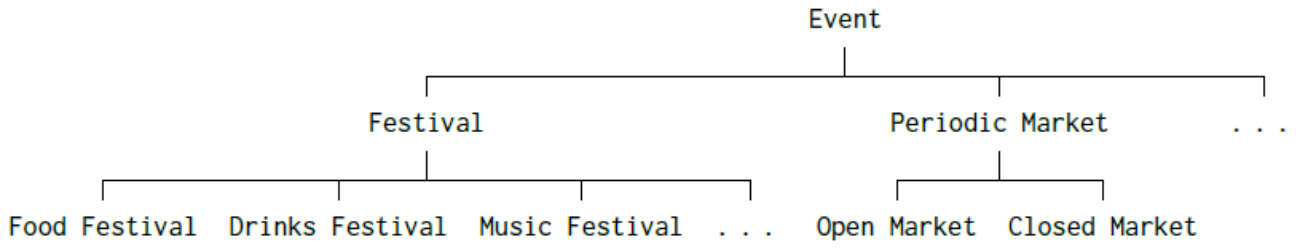


Fig. 10. Example class hierarchy for Event

Table 2. Definitions for sub-class, sub-property, domain and range features in semantic schemata

Feature	Definition	Condition	Example
SUBCLASS	c —subc. of→ d	x —type→ c implies x —type→ d	City —subc. of→ Place
SUBPROPERTY	p —subp. of→ q	x —p→ y implies x —q→ y	venue —subp. of→ location
DOMAIN	p —domain→ c	x —p→ y implies x —type→ c	venue —domain→ Event
RANGE	p —range→ c	x —p→ y implies y —type→ c	venue —range→ Venue

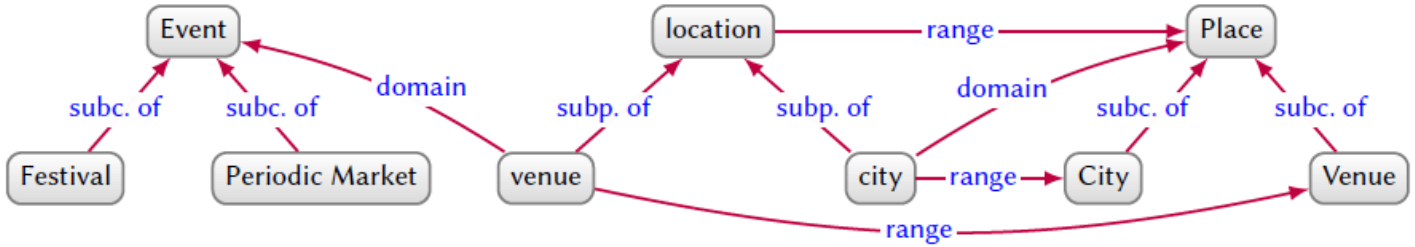


Fig. 11. Example schema graph describing sub-classes, sub-properties, domains, and ranges

Definition of higher-level terms used in the graph (*vocabulary*, *taxonomy, …*)

- Special label **type** allows to type nodes with higher-level terms
- Type hierarchies
- Property hiererachies
- Property domain and range definitions
- More complex logical axioms

✓ One can **reason on a graph** wrt a semantic schema

example for W3C RDF:

✓ RDF Schema (RDFS)

✓ Web Ontology Language (OWL)

Open vs Closed world assumption

Schema 2/3: *validation* type

Enables to verify that a graph is **complete**
That important data is **explicit**

✓ One may **validate a graph** wrt a validation schema

exemple for W3C RDF:
✓ SHACL (Shapes Constraint Language)
✓ ShEx (Shape Expressions)

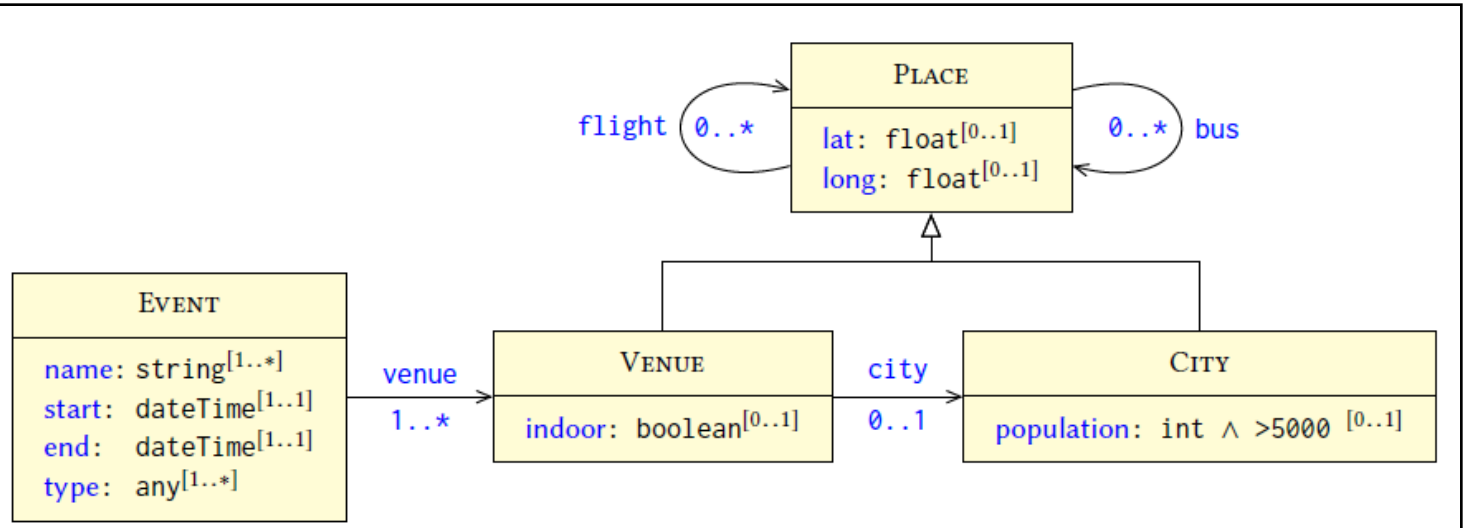
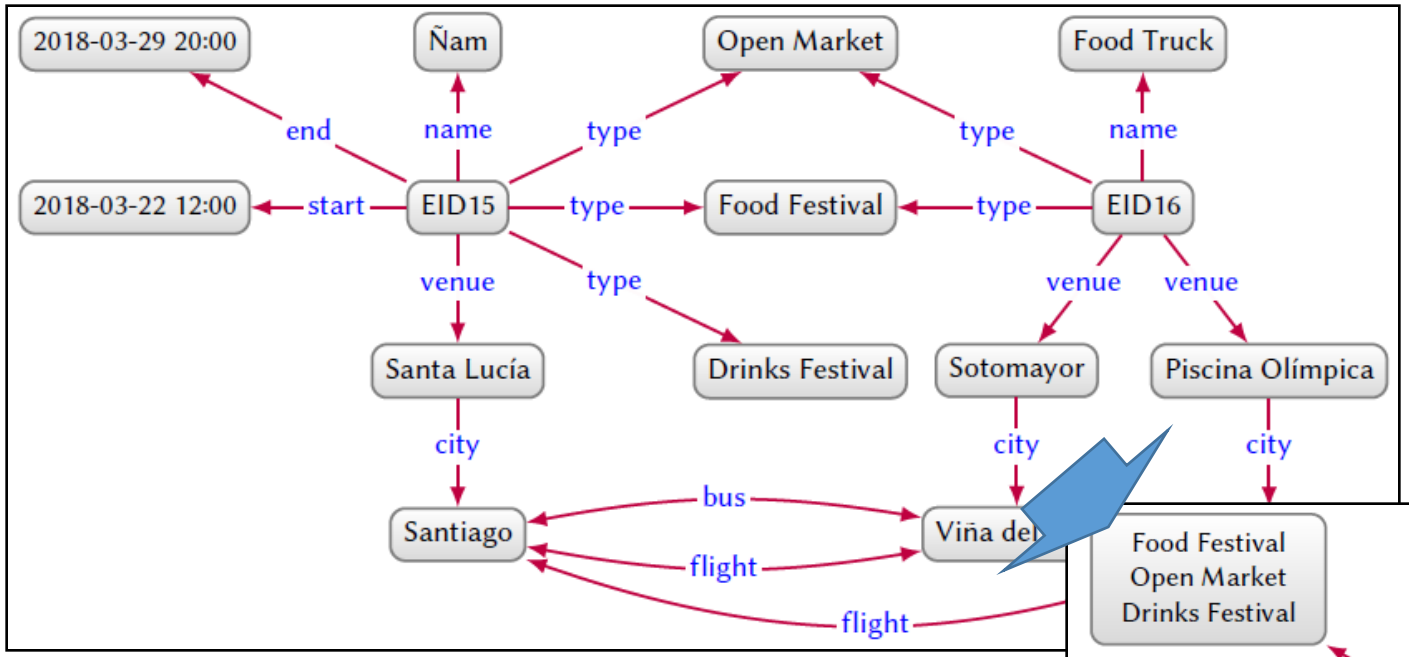


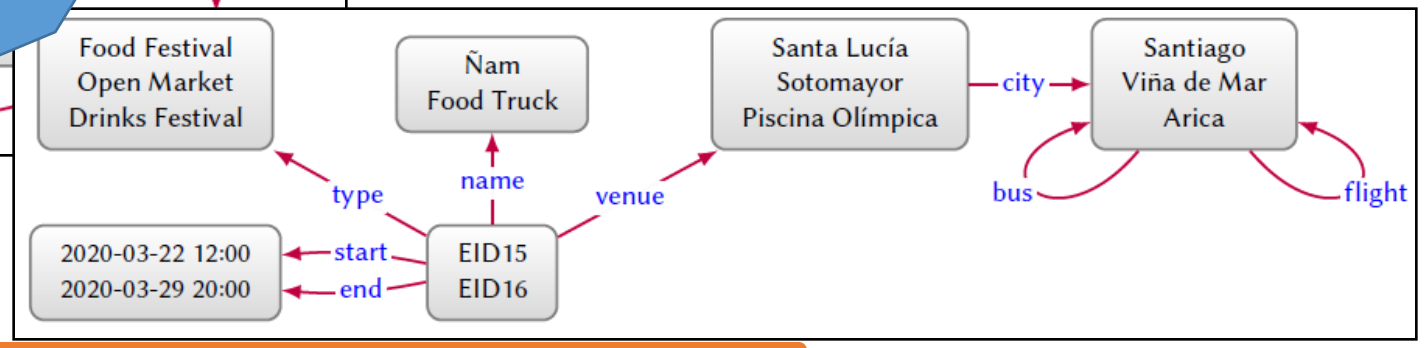
Fig. 12. Example shapes graph depicted as a UML-like diagram

Schema 2/3: *emerging* type



One tries to merge nodes while preserving arcs:

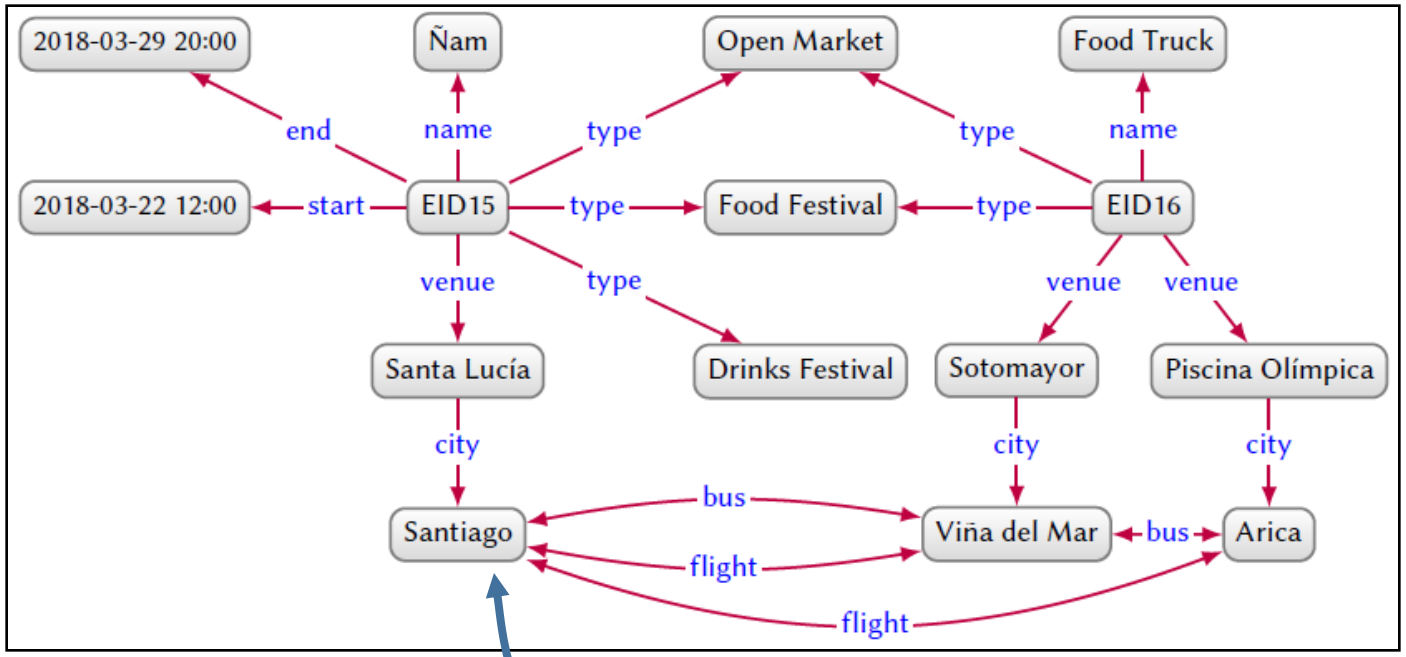
✓ One may **summarize a graph** un a **quotient graph**



Definition B.28 (Quotient graph). Given a directed-edge labelled graph $G = (V, E, L)$, a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, L)$ is a *quotient graph* of G if and only if:

- $\mathcal{V}$ is a partition of V without the empty set, i.e., $\mathcal{V} \subseteq (2^V - \emptyset)$, $V = \bigcup_{U \in \mathcal{V}} U$, and for all $U \in \mathcal{V}, W \in \mathcal{V}$, it holds that $U = W$ or $U \cap W = \emptyset$; *and*
- $\mathcal{E} = \{(U, l, W) \mid U \in \mathcal{V}, W \in \mathcal{V}$ and there exist $u \in U, w \in W$ such that $(u, l, w) \in E\}$.

Identity



Santiago de Chile ?
Santiago de Cuba ?
Santiago de Compostela ?
Santiago indie rock band ?

Need to **deambiguise** nodes

- for graphs fusion
- for graphs exchange

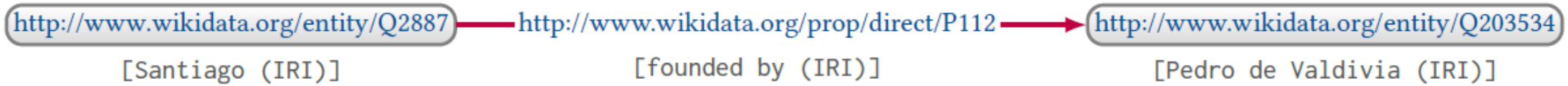
Identity: persistent identifiers

✓ Use persistent identifiers to label nodes

- *Digital Object Identifiers (DOIs)* for articles
- *ORCID iDs* for authors
- *Alpha-2 codes* for countries
- ENTSO-E Energy Identification Coding (EIC) scheme
  -> Party, Area, Accounting Point, Tie Line, Location, Resource Object, Substation
- …

On the Semantic Web:

✓ Use *Internationalised Resource Identifiers (IRIs)*



Optional: shorten using “namespaces”
wd: = http://www.wikidata.org/entity/ , wdt := http://www.wikidata.org/prop/direct/

Identity
links to external identifiers

It is possible to use IRIs defined in an external knowledge graph

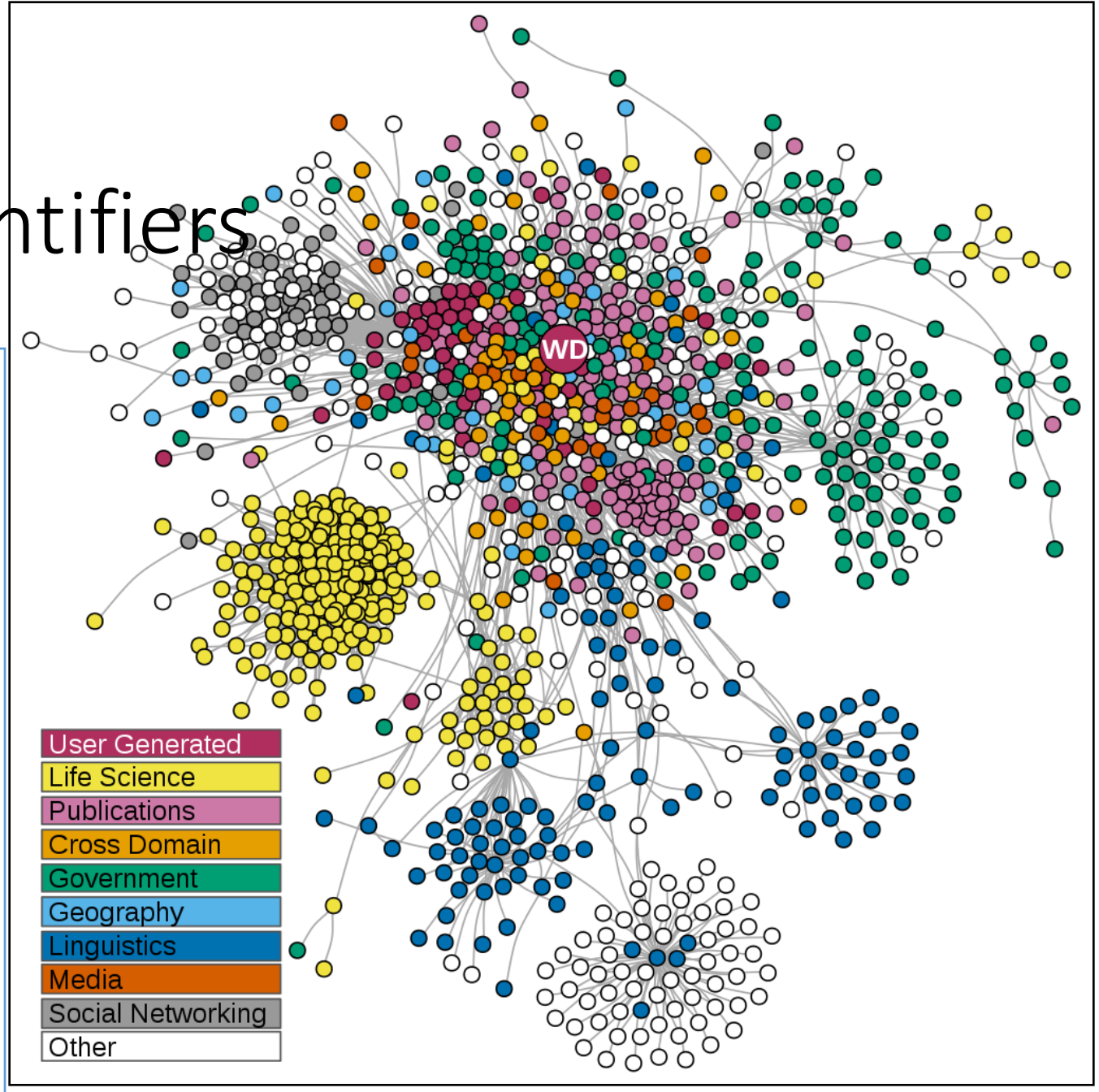
- to use a reference identification system
- to expand a knowledge graph

anyone can say anything about anything

The co-reference of two nodes can be determined using a semantic schema.
examples:

- ✓ W3C OWL owl:FunctionalProperty type
- ✓ the W3C OWL owl:sameAs relationship

chile:Santiago - owl:sameAs → geo:SantiagoDeChile



Identity: data types

No persistent identifier for dates, integers, ...

- ✓ but a **data type**
- ✓ a persistent identifier for the data type

In the Semantic Web:

- ✓ RDF uses XML Schema Datatypes (XSD)
- ✓ A **data type** *d* is identified by an IRI, and associates a valid character string (in *d's* **lexical space**) to a value (in *d's* **value space**)
- ✓ Nodes can be labeled by a **literal**: a pair < string , data type IRI >.
- ✓ A node labeled with a literal has an outgoing degree equal to zero.

- ➢ example data types: xsd:string, xsd:integer, xsd:decimal, xsd:boolean
- ➢ example of a literal `"2020-03-29T20:00:00"^^xsd:dateTime`

In Neo4j:

- ➢ example data types: numbers, strings, booleans, spatial points, and temporal values.

Identity: lexicalization

It is irrelevant whether the identifier can be interpreted by a human.

example wikidata uses opaque identifiers wd:Q2887

In practice, it is useful to include text elements in the knowledge graph.

✓ we can use a small consensual set of properties

wd:Q2887 - rdfs:label → "Santiago"

wd:Q2887 - skos:altLabel → "Santiago de Chile"

wd:Q2887 - rdfs:comment → "Santiago is the capital of Chile"

In the Semantic Web:

✓ A **literal** can have a language label

➢ example of literal with language label

chile:City - rdfs:label → "City"@en

chile:City - rdfs:label → "Ciudad"@es

Identity: existential nodes

It**'s handy** to have the option of **not labeling** certain nodes

We use **existential nodes**, without etiquette.

example: to model a list with a directed labeled graph model (with a chained list model), we don't want to identify each sub-list.

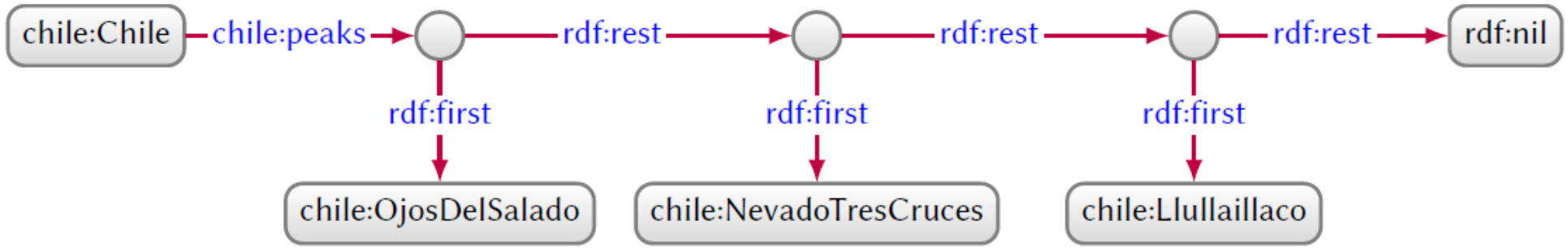


Fig. 16. RDF list representing the three largest peaks of Chile, in order

It's practical in theory, but it complicates the algorithms a lot.

example: deciding whether two graphs are isomorphic modulo the anonymous nodes.

Context

All knowledge is true in a certain context

- temporal context

example: Santiago existed as a town since 1541, etc.

- geographical context

- context of origin

example: *the Wikipedia France version on January 4, 2021*

- agent context

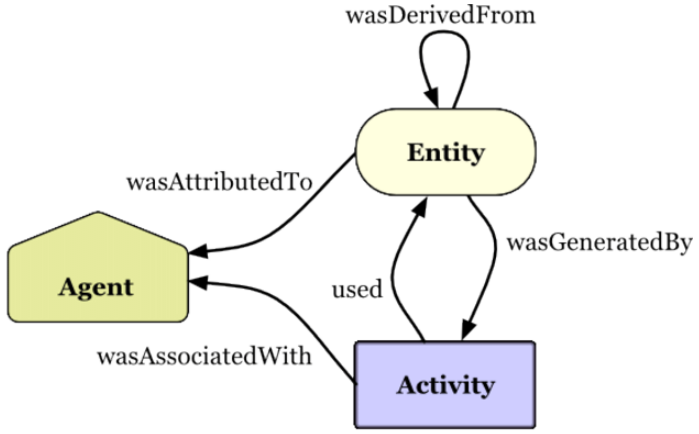
example: agent x believes that ... , agent a knows that ...

Context: direct representation

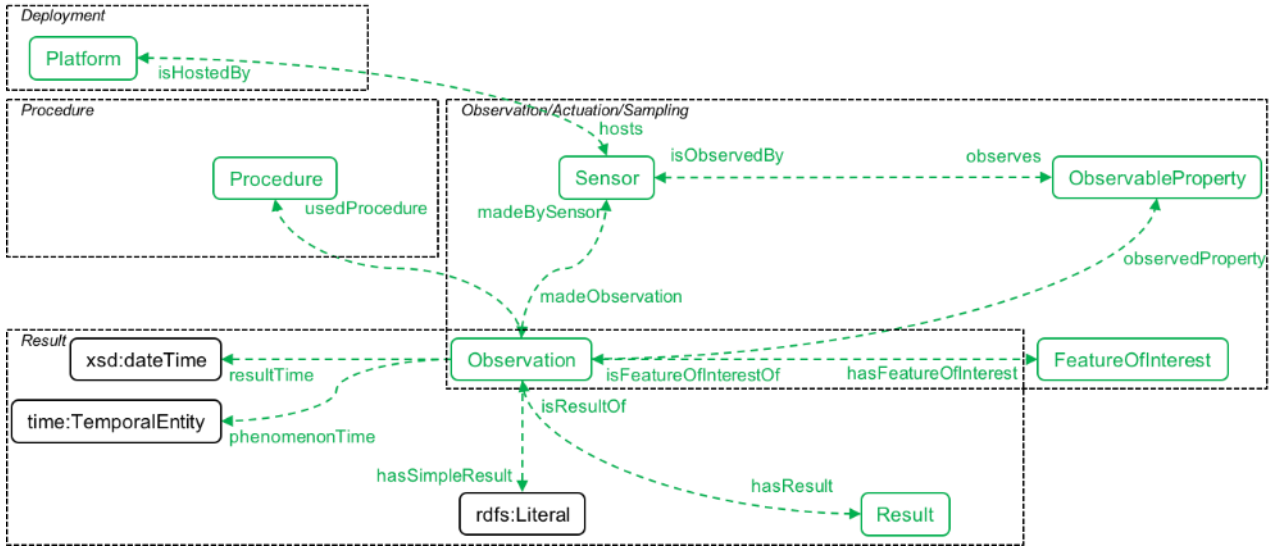
Use of consensual vocabulary to represent context

for example:

- W3C PROV Data Model, to describe where entities come from and how they are derived from other entities, generated and/or used by activities, assigned to agents.
- W3C Semantic Sensor Network ontology, to provide the context for assigning a value to a property of a thing of interest by a sensor



source: http://www.w3.org/TR/prov-primer/



source: https://www.w3.org/TR/vocab-ssn/

Context: reification

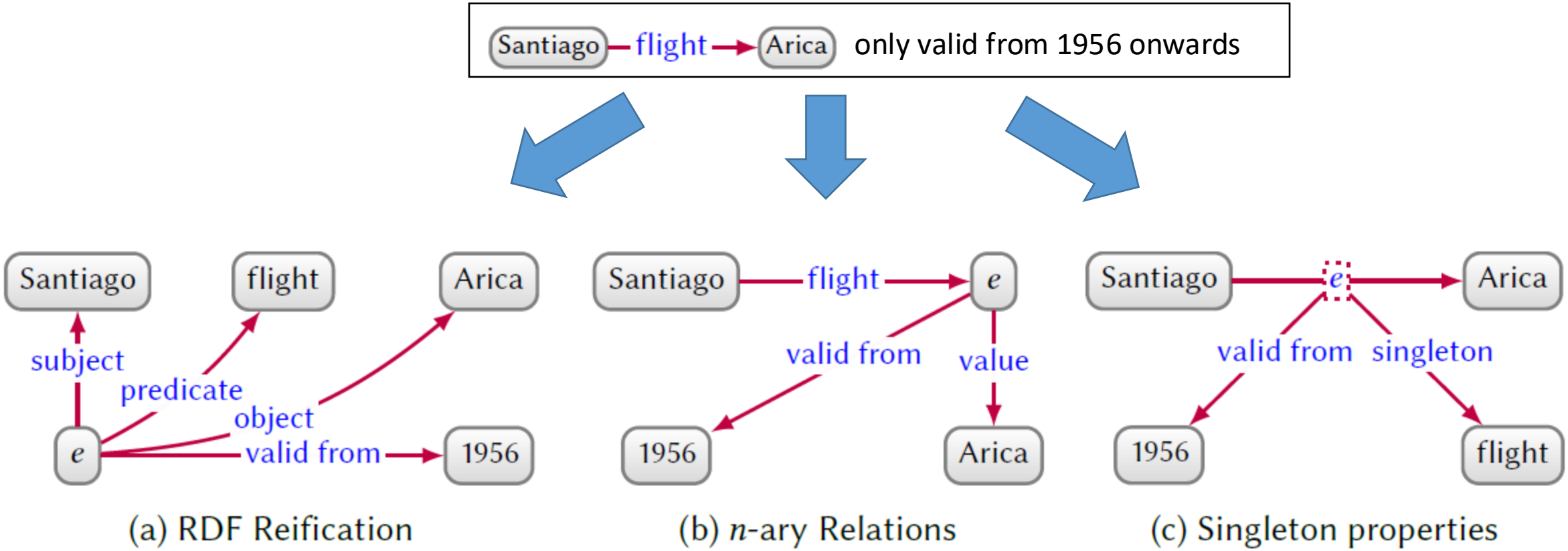


Fig. 17. Three representations of temporal context on an edge in a directed-edge labelled graph

Context: representation by greater arity

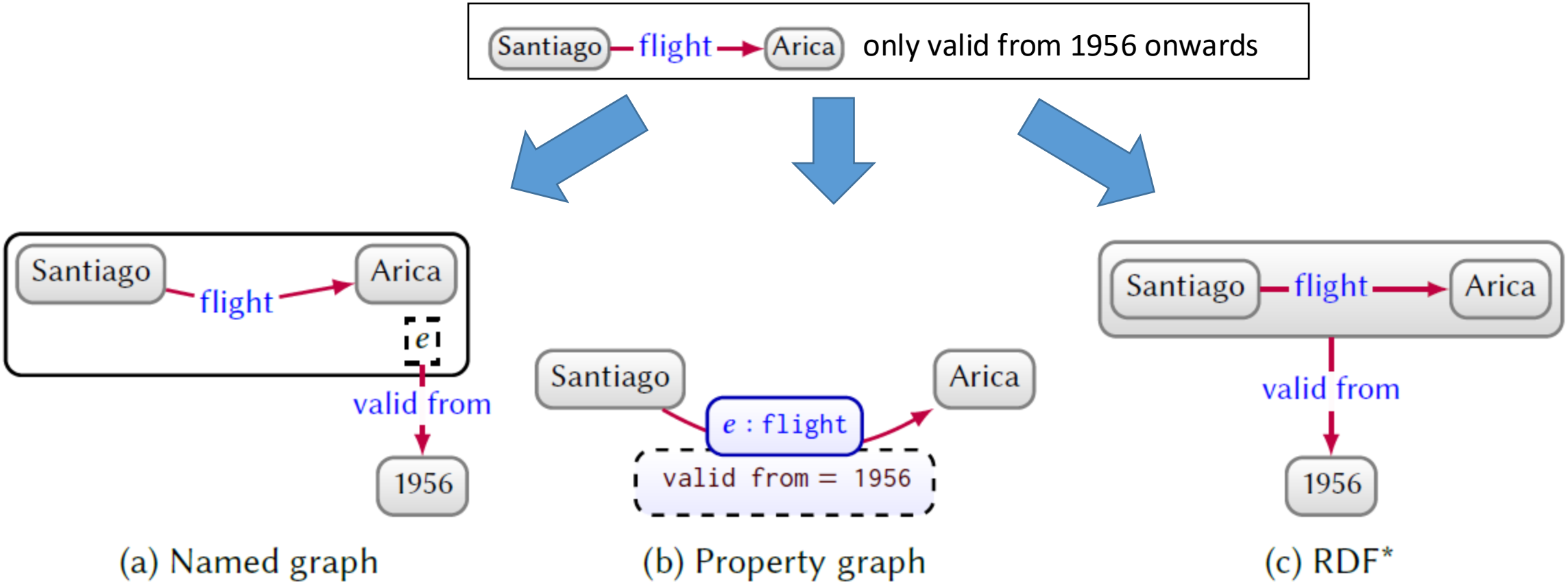
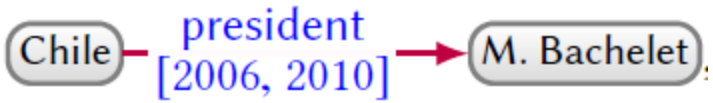
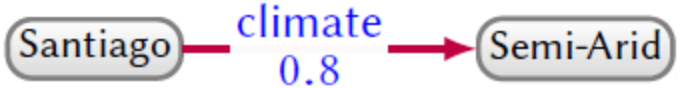


Fig. 18. Three higher-arity representations of temporal context on an edge

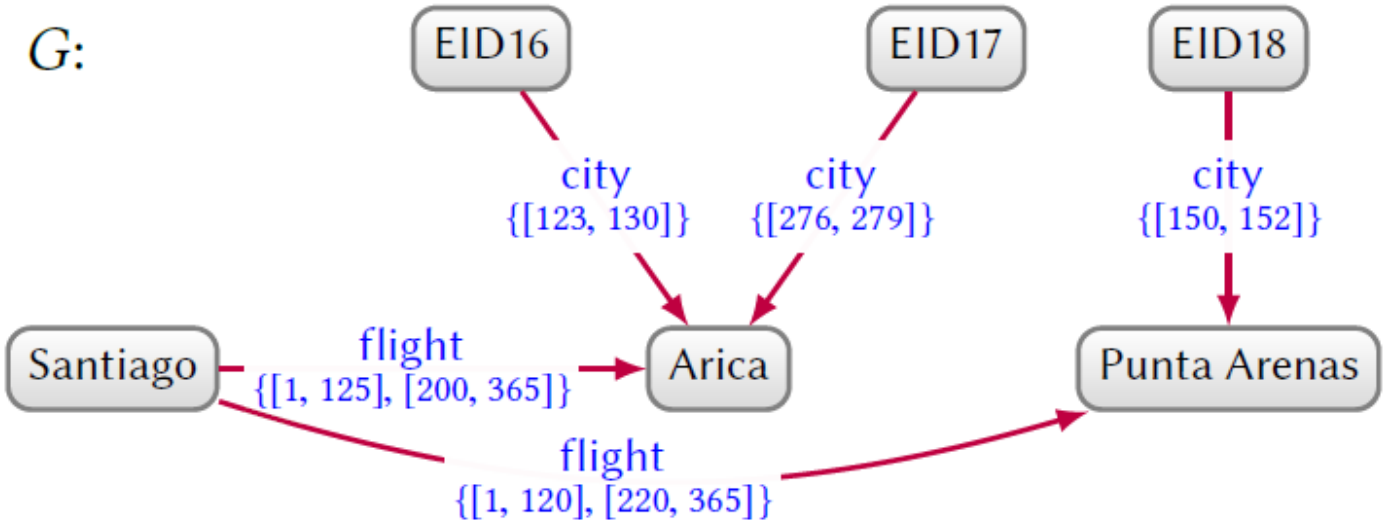
Context: annotations

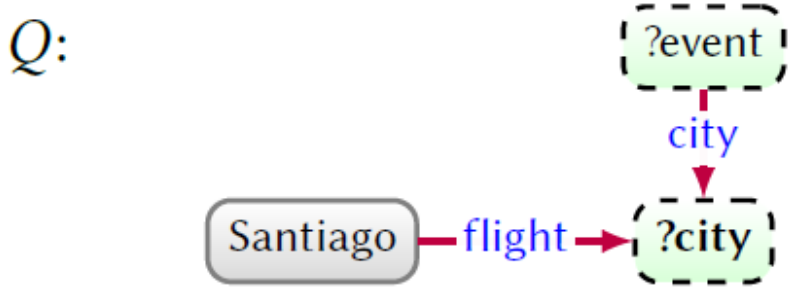
Use of a mathematical definition for the context domain of arc validity

- example: Temporal RDF 

- example: Fuzzy RDF 

- example: Annotated RDF (for domains that can be defined by an algebraic semi-ring structure)

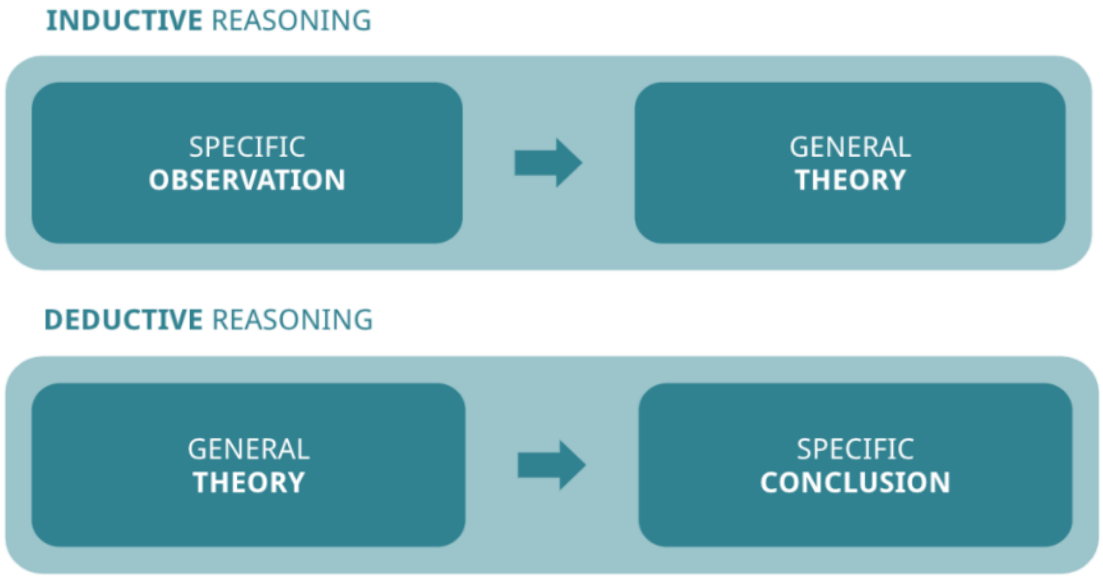




$Q(G)$:

?city	context
Arica	{[123, 125], [276, 279]}

Inductive knowledge vs. deductive knowledge



source: https://nesslabs.com/inductive-deductive-reasoning

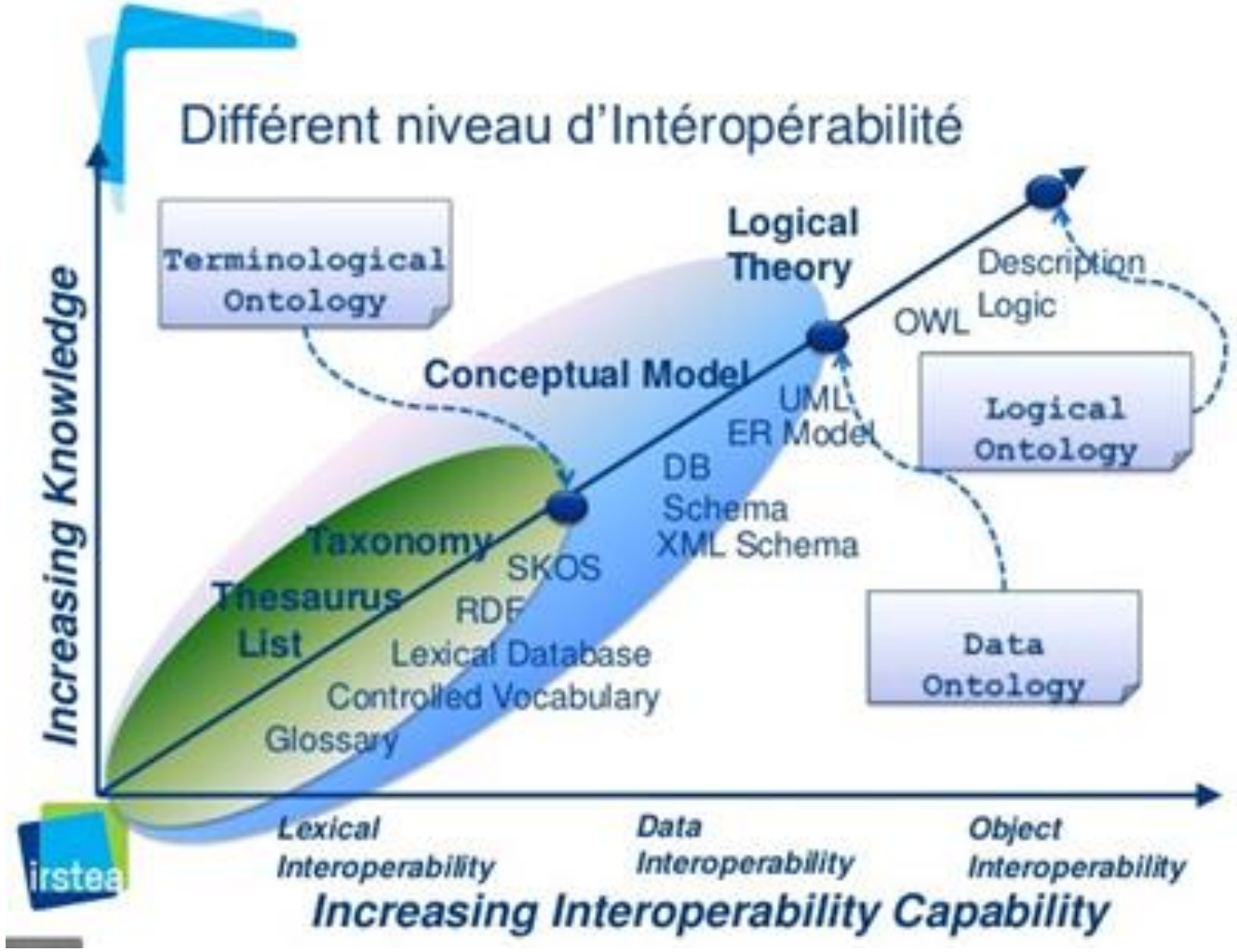
Ontologies

Explicit, formal specification of a shared conceptualization

- A concrete, formal representation of what a term means in the context of its use

In the Semantic Web

- The W3C standard OWL (Web Ontology Language) is used to define ontologies.



source: https://fr.slideshare.net/croussey/skos-transformation

Web Ontology Language

Table 3. Ontology features for individuals

Feature	Axiom	Condition	Example
Assertion	$x \xrightarrow{y} z$	$x \overset{y}{\Rightarrow} z$	Chile —capital→ Santiago
Negation	n —type→ Neg; n —sub→ x; n —pre→ y; n —obj→ z	not $x \overset{y}{\Rightarrow} z$	n —type→ Neg; n —sub→ Chile; n —pre→ capital; n —obj→ Arica
Same As	x_1 —same as→ x_2	$x_1 = x_2$	Región V —same as→ Región de Valparaíso
Different From	x_1 —diff. from→ x_2	$x_1 \neq x_2$	Valparaíso —diff. from→ Región de Valparaíso

Table 4. Ontology features for property axioms

Feature	Axiom	Condition (for all x_*, y_*, z_*)	Example
Subproperty	p —subp. of→ q	$x \overset{p}{\Rightarrow} y$ implies $x \overset{q}{\Rightarrow} y$	venue —subp. of→ location
Domain	p —domain→ c	$x \overset{p}{\Rightarrow} y$ implies $x \overset{\text{type}}{\Rightarrow} c$	venue —domain→ Event
Range	p —range→ c	$x \overset{p}{\Rightarrow} y$ implies $y \overset{\text{type}}{\Rightarrow} c$	venue —range→ Venue
Equivalence	p —equiv. p.→ q	$x \overset{p}{\Rightarrow} y$ iff $x \overset{q}{\Rightarrow} y$	start —equiv. p.→ begins
Inverse	p —inv. of→ q	$x \overset{p}{\Rightarrow} y$ iff $y \overset{q}{\Rightarrow} x$	venue —inv. of→ hosts
Disjoint	p —disj. p.→ q	not $x \overset{q}{\underset{p}{\rightrightarrows}} y$	venue —disj. p.→ hosts
Transitive	p —type→ Transitive	$x \overset{p}{\Rightarrow} y \overset{p}{\Rightarrow} z$ implies $x \overset{p}{\Rightarrow} z$	part of —type→ Transitive
Symmetric	p —type→ Symmetric	$x \overset{p}{\Rightarrow} y$ iff $y \overset{p}{\Rightarrow} x$	nearby —type→ Symmetric
Asymmetric	p —type→ Asymmetric	not $x \overset{p}{\underset{p}{\leftrightarrows}} y$	capital —type→ Asymmetric
Reflexive	p —type→ Reflexive	$x \overset{p}{\Rightarrow} x$	part of —type→ Reflexive
Irreflexive	p —type→ Irreflexive	not $x \overset{p}{\Rightarrow} x$	flight —type→ Irreflexive
Functional	p —type→ Functional	$y_1 \overset{p}{\Leftarrow} x \overset{p}{\Rightarrow} y_2$ implies $y_1 = y_2$	population —type→ Functional
Inv. Functional	p —type→ Inv. Functional	$x_1 \overset{p}{\Rightarrow} y \overset{p}{\Leftarrow} x_2$ implies $x_1 = x_2$	capital —type→ Inv. Functional

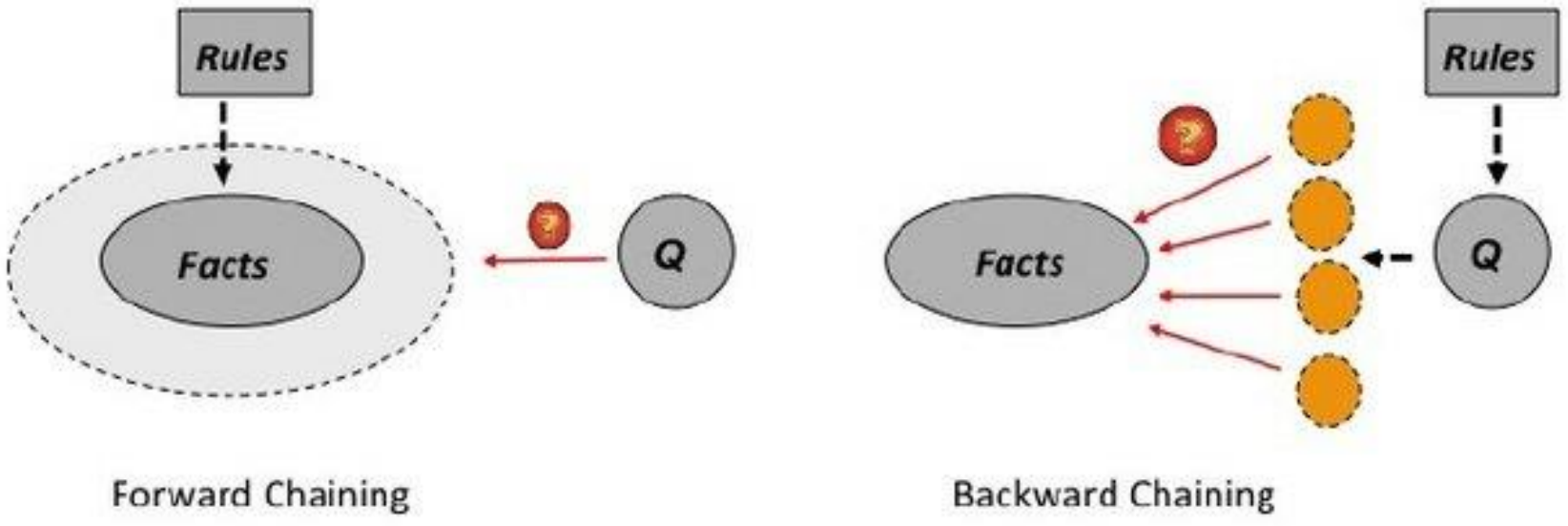
Table 5. Ontology features for class axioms and definitions

Feature	Axiom	Condition (for all x_*, y_*, z_*)	Example
Subclass	c —subc. of→ d	$x \overset{\text{type}}{\Rightarrow} c$ implies $x \overset{\text{type}}{\Rightarrow} d$	City —subc. of→ Place
Equivalence	c —equiv. c.→ d	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \overset{\text{type}}{\Rightarrow} d$	Human —equiv. c.→ Person
Disjoint	c —disj. c.→ d	not $c \overset{\text{type}}{\Leftarrow} x \overset{\text{type}}{\Rightarrow} d$	City —disj. c.→ Region
Complement	c —comp.→ d	$x \overset{\text{type}}{\Rightarrow} c$ iff not $x \overset{\text{type}}{\Rightarrow} d$	Dead —comp.→ Alive
Union	c —union→ $d_1, \ldots, d_n$	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \overset{\text{type}}{\Rightarrow} d_1$ or … or $x \overset{\text{type}}{\Rightarrow} d_n$	Flight —union→ DomesticFlight, InternationalFlight
Intersection	c —inter.→ $d_1, \ldots, d_n$	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \overset{\text{type}}{\Rightarrow} d_1, \ldots, x \overset{\text{type}}{\Rightarrow} d_n$	SelfDrivingTaxi —inter.→ Taxi, SelfDriving
Enumeration	c —one of→ $x_1, \ldots, x_n$	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \in \{x_1, \ldots, x_n\}$	EUState —one of→ Austria, …, Sweden
Some Values	c —prop some→ d	$x \overset{\text{type}}{\Rightarrow} c$ iff there exists a such that $x \overset{p}{\Rightarrow} a \overset{\text{type}}{\Rightarrow} d$	EUCitizen —prop some→ nationality, EUState
All Values	c —prop all→ d	$x \overset{\text{type}}{\Rightarrow} c$ iff for all a with $x \overset{p}{\Rightarrow} a$ it holds that $a \overset{\text{type}}{\Rightarrow} d$	Weightless —prop all→ has part, Weightless
Has Value	c —prop value→ y	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \overset{p}{\Rightarrow} y$	ChileanCitizen —prop value→ nationality, Chile
Has Self	c —prop self→ true	$x \overset{\text{type}}{\Rightarrow} c$ iff $x \overset{p}{\Rightarrow} x$	SelfDriving —prop self→ driver, true
Cardinality $\star \in \{=, \leq, \geq\}$	c —prop $\star$→ n	$x \overset{\text{type}}{\Rightarrow} c$ iff $\#\{a \mid x \overset{p}{\Rightarrow} a\} \star n$	Polyglot —prop ≥→ fluent, 2
Qualified Cardinality $\star \in \{=, \leq, \geq\}$	c —prop class→ d; c —prop $\star$→ n	$x \overset{\text{type}}{\Rightarrow} c$ iff $\#\{a \mid x \overset{p}{\Rightarrow} a \overset{\text{type}}{\Rightarrow} d\} \star n$	BinaryStarSystem —prop class→ Star; —prop =→ 2; —prop →body

Rule-based reasoning

Simplest forms of rule-based reasoning

- front chain
- rear chain

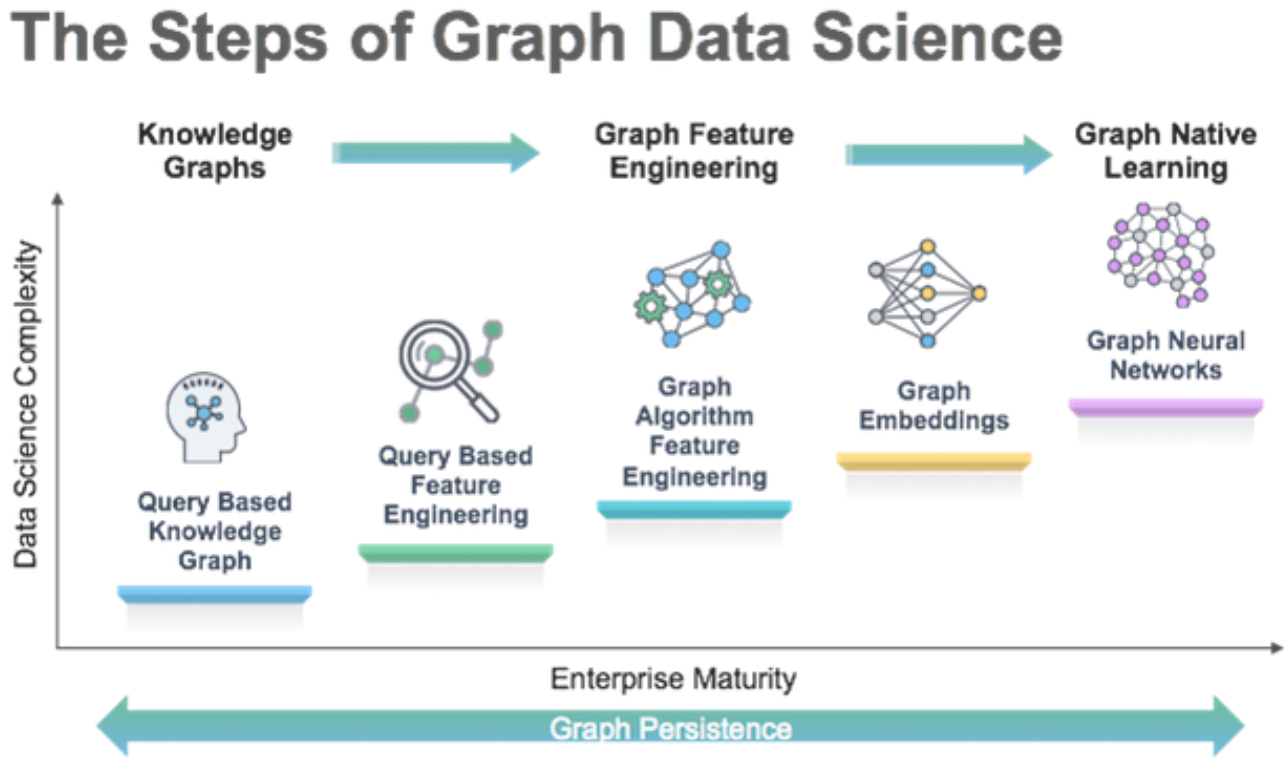


source: König, Mélanie & Leclere, Michel & Mugnier, Marie-Laure & Thomazo, Michaël (2013). Sound, Complete and Minimal UCQ-Rewriting for Existential Rules. Semantic Web. 6. 10.3233/SW-140153.

Knowledge induction

- **Deduction is infallible, but complex**
  (both to develop the ontology/rules and to reason)

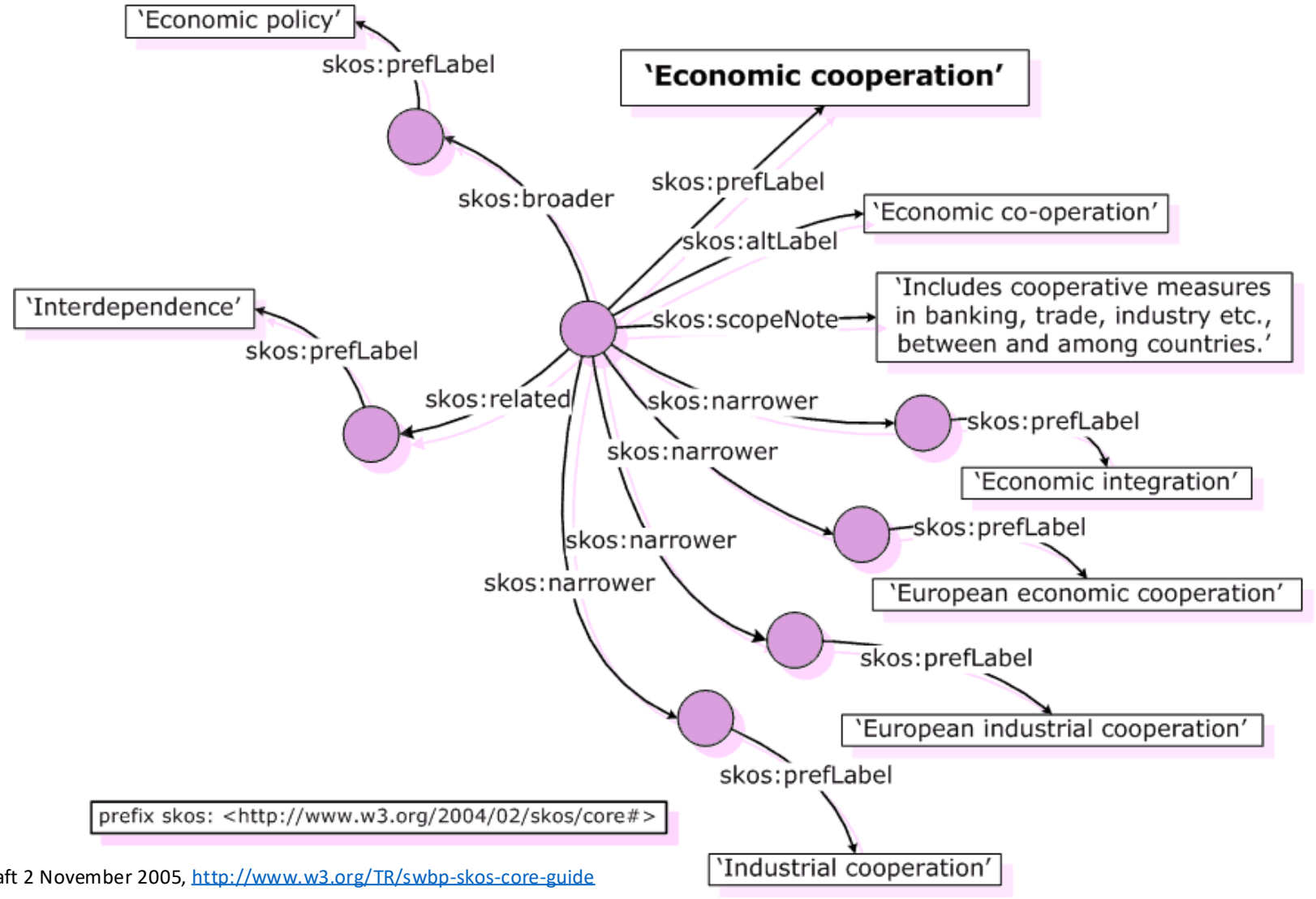
- **Induction is fallible, but can provide valuable insights**



source: https://neo4j.com/blog/deepwalk-implementing-graph-embeddings-in-neo4j/

Vocabularies, taxonomies, ontologies

SKOS Simple Knowledge Organization System



SKOS Core Guide, W3C Working Draft 2 November 2005, http://www.w3.org/TR/swbp-skos-core-guide

Examples of thesauri, taxonomies, …

UNESCO Thesaurus

The UNESCO Thesaurus is **a controlled and structured list of terms used in subject analysis and retrieval of documents and publications** in the fields of education, culture, natural sciences, social and human sciences, communication and information.

https://skos.um.es › unescothes

UNESCO Thesaurus - SKOS

GEMET General Multilingual Environmental Thesaurus

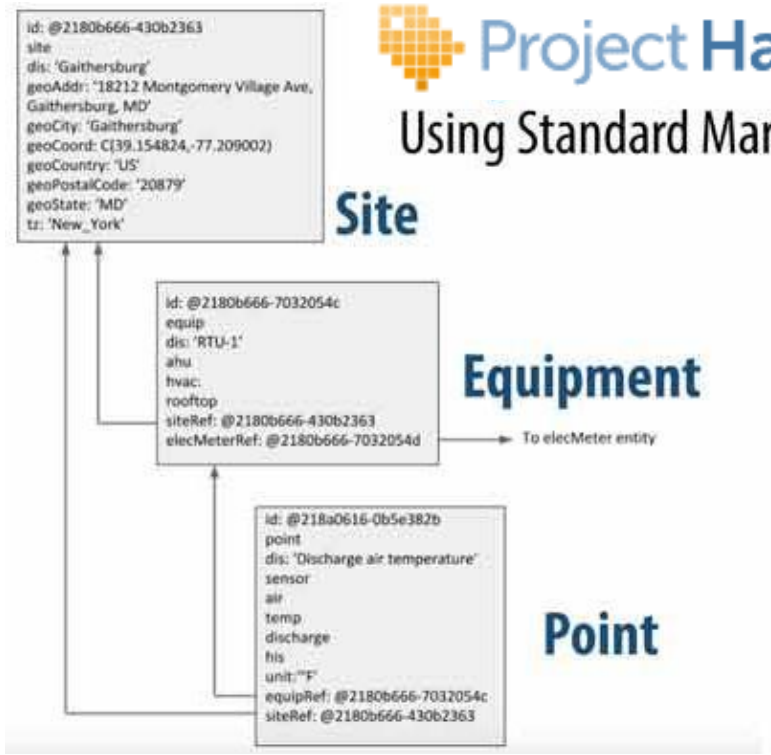
https://www.eionet.europa.eu/gemet/en/webservices/

nasa/**dictionaries**

A collection of NASA "dictionaries", including thesauri, taxonomies and ontologies.

2 Contributors · 2 Issues · 51 Stars · 31 Forks

https://github.com/nasa/dictionaries

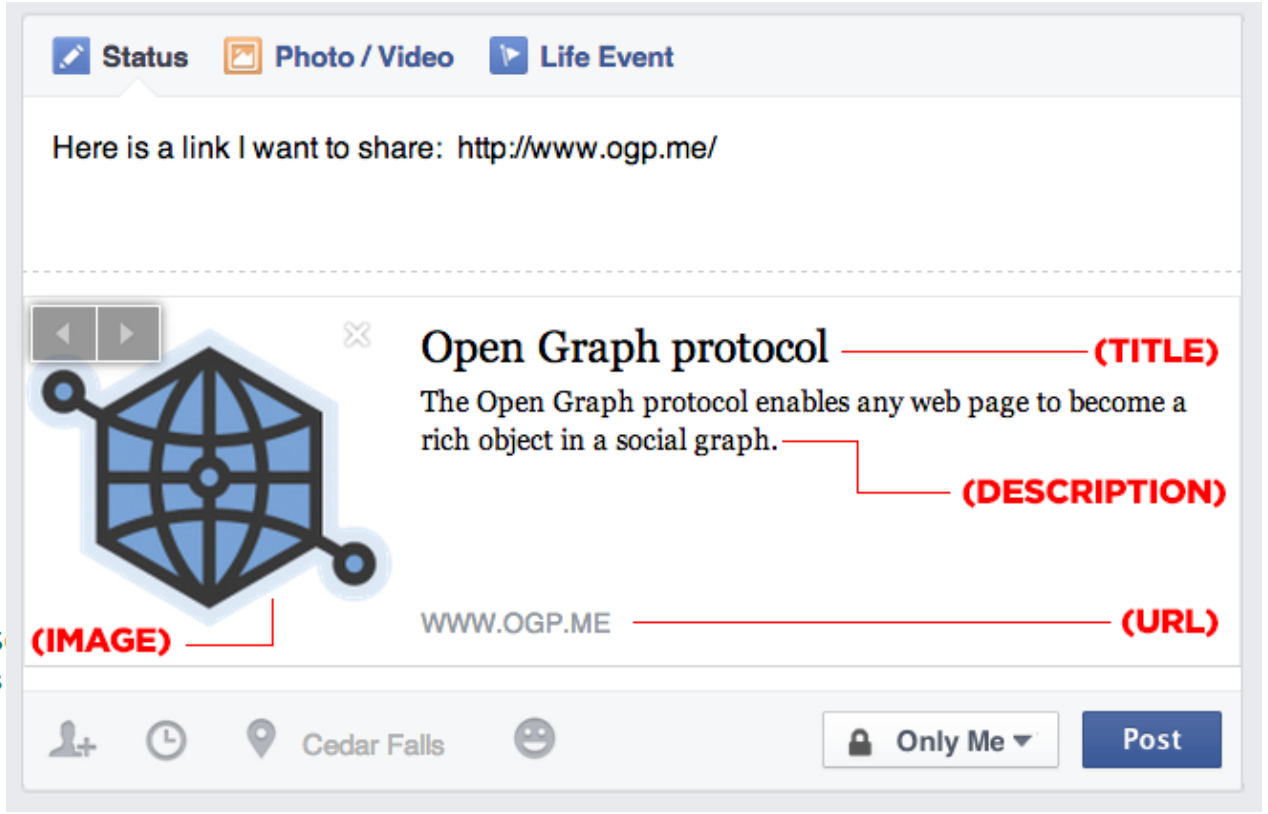


https://project-haystack.org/

List at https://www.w3.org/2001/sw/wiki/SKOS/Datasets

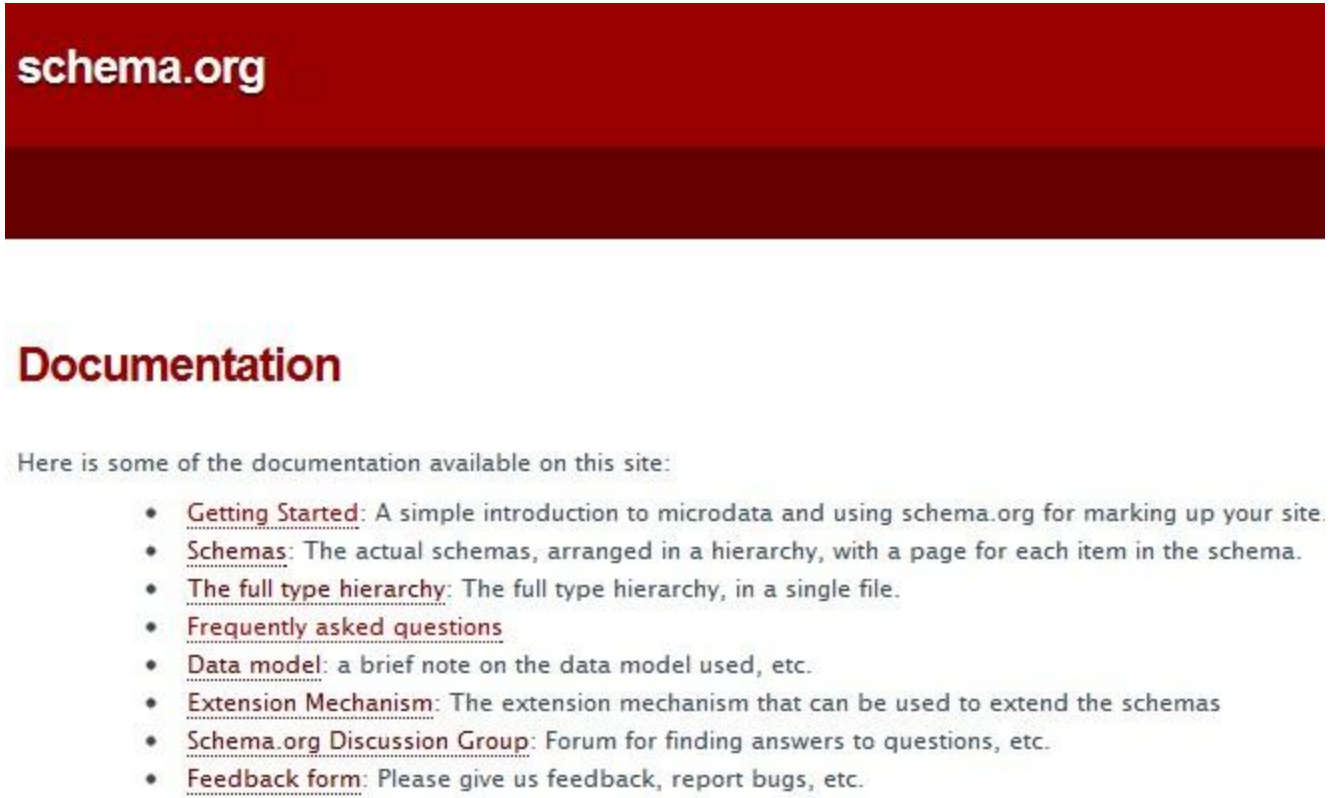
Examples of RDF vocabularies

```
<html prefix="og: https://ogp.me/ns#">
<head>
<title>The Rock (1996)</title>
<meta property="og:title" content="The Rock" />
<meta property="og:type" content="video.movie" />
<meta property="og:url" content="https://www.imdb.com/title/tt01175
<meta property="og:image" content="https://ia.media-imdb.com/images
...
</head>
...
</html>
```



Open Graph Protocol https://ogp.me/ (Facebook, 2010), test with https://developers.facebook.com/tools/debug/

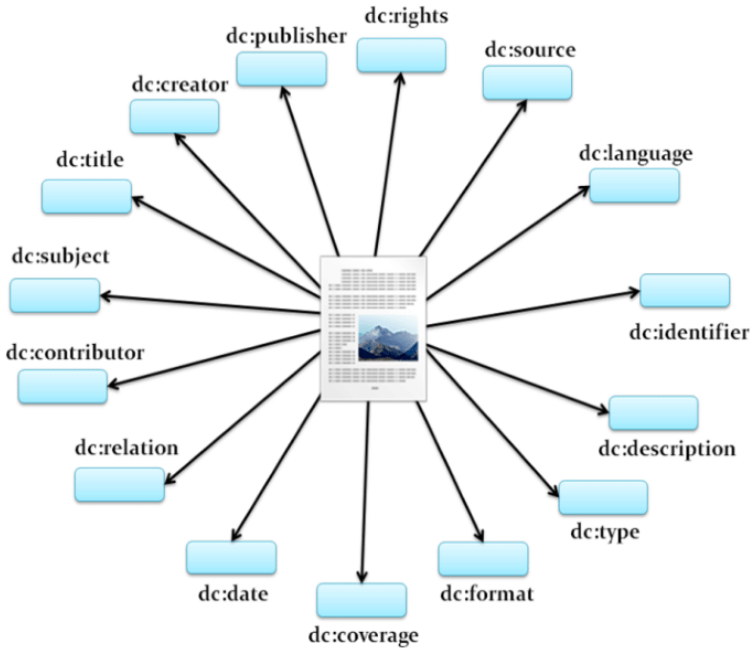
Examples of RDF vocabularies



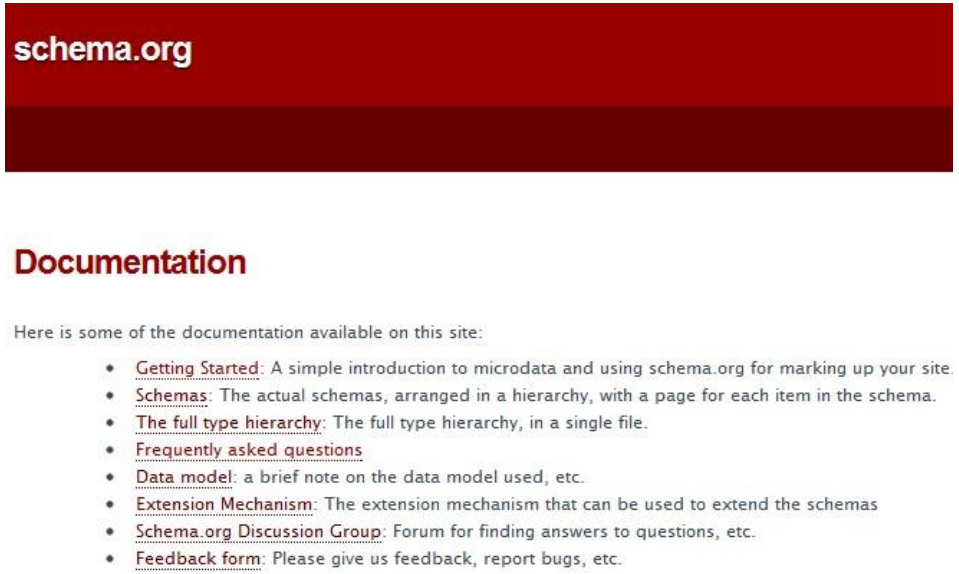
Over 10 million sites use Schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences.

Founded by Google, Microsoft, Yahoo and Yandex, 2011
Test with https://validator.schema.org/ or https://developers.google.com/search/docs/advanced/structured-data

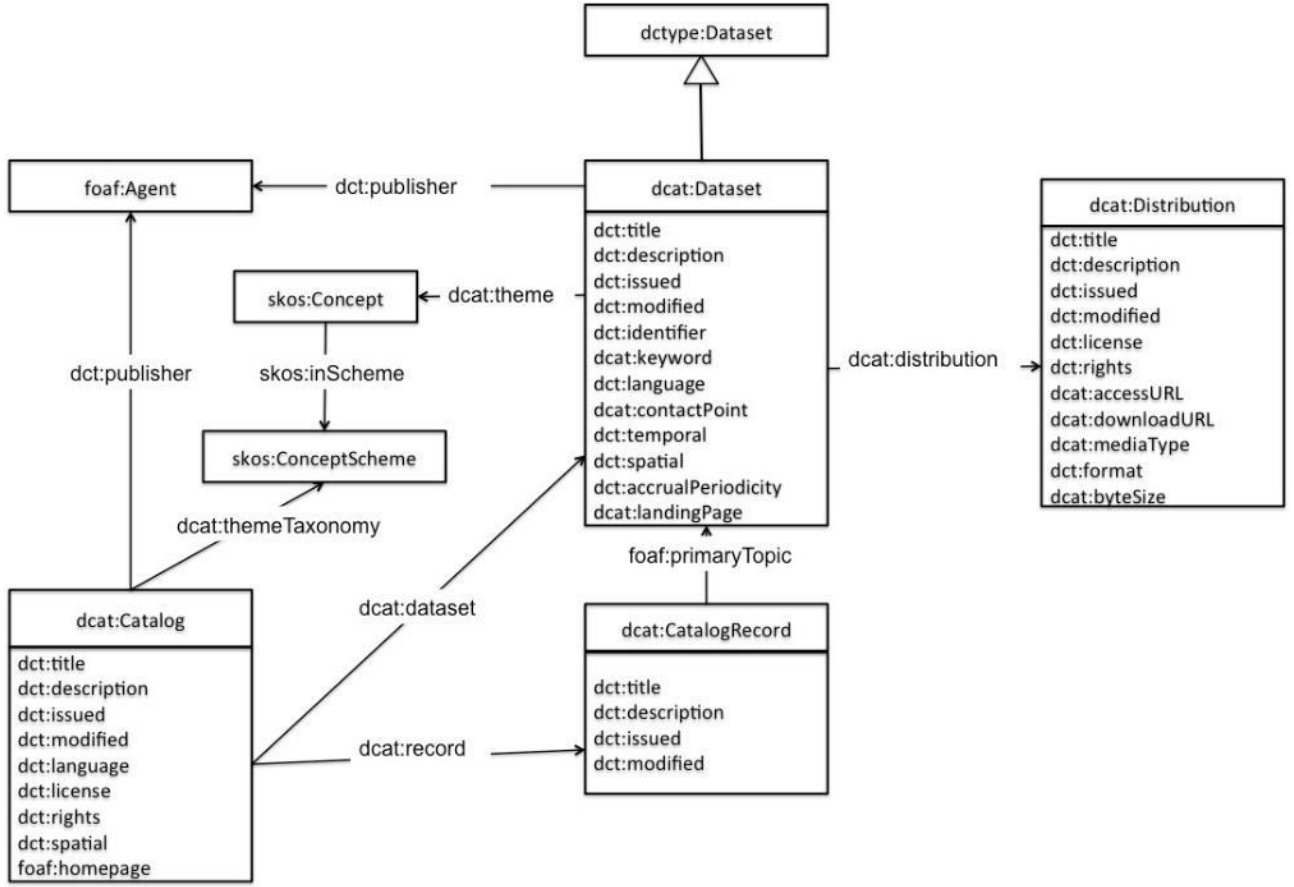
Examples of RDF vocabularies for metadata



Dublin Core



schema.org



Data Catalog Vocabulary (DCAT) - Version 2
W3C Recommendation 04 February 2020

Examples of RDF ontologies for the Energy domain

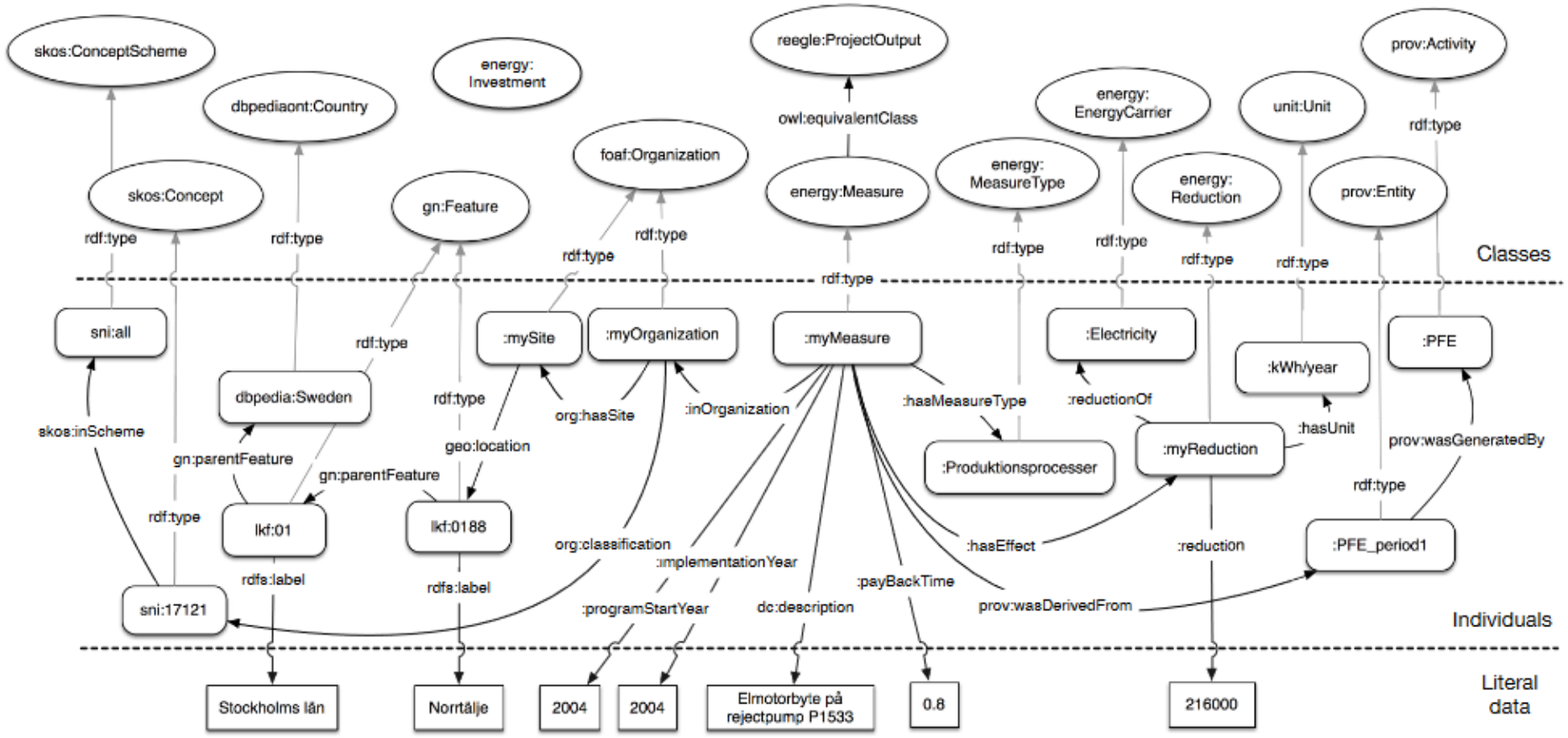


Fig. 2. Example data illustrating one measure originating from the PFE project as represented in our dataset, using our vocabulary (energy:). Entries given without namespace prefix are defined in the relevant version of the local data namespace.

Blomqvist, E., Thollanderb, P., Keskisärkkä, R., & Paramonovab, S. (2014). Energy efficiency measures as linked open data. IOS Press, 1, 1-5.

Examples of RDF ontologies for the Energy domain

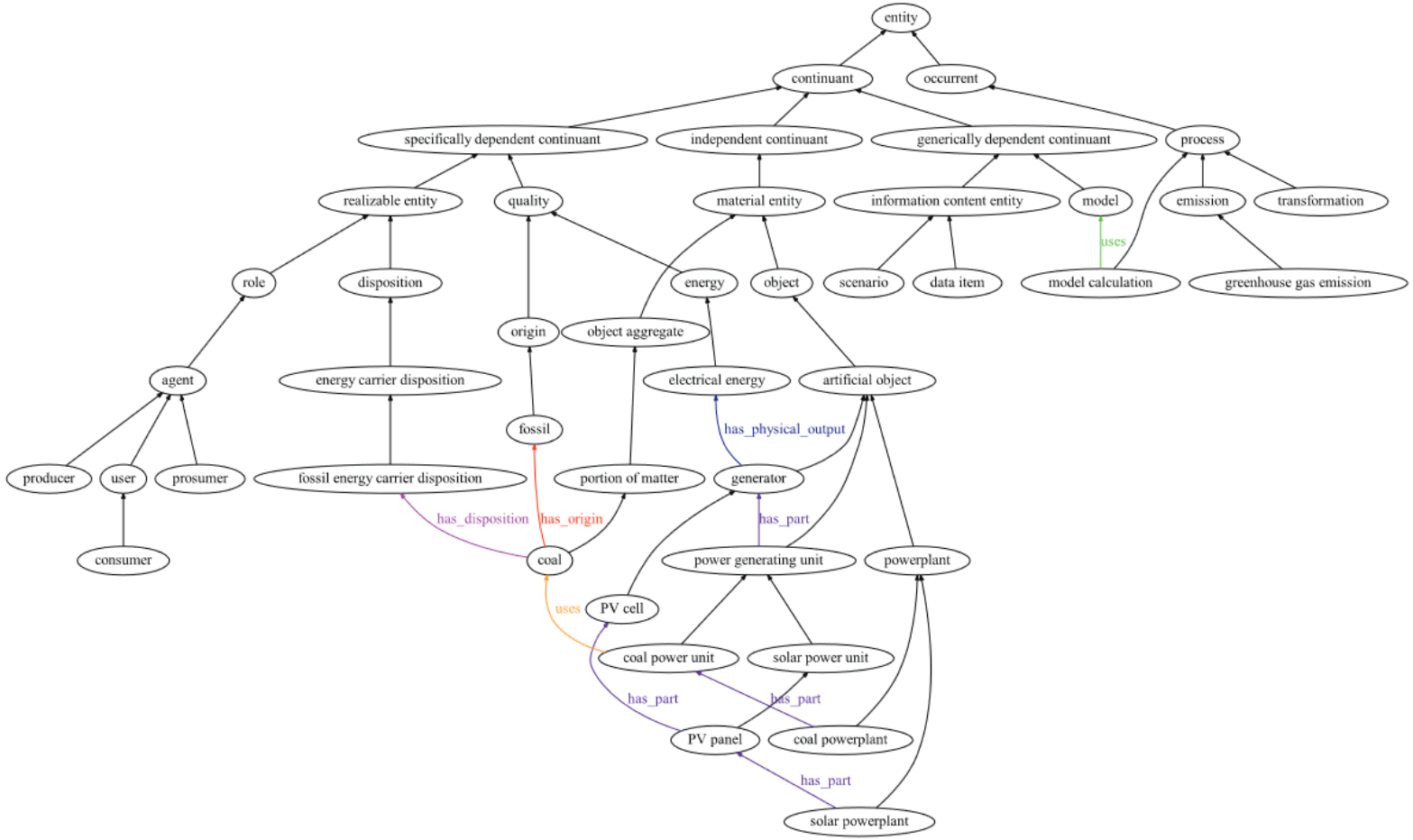
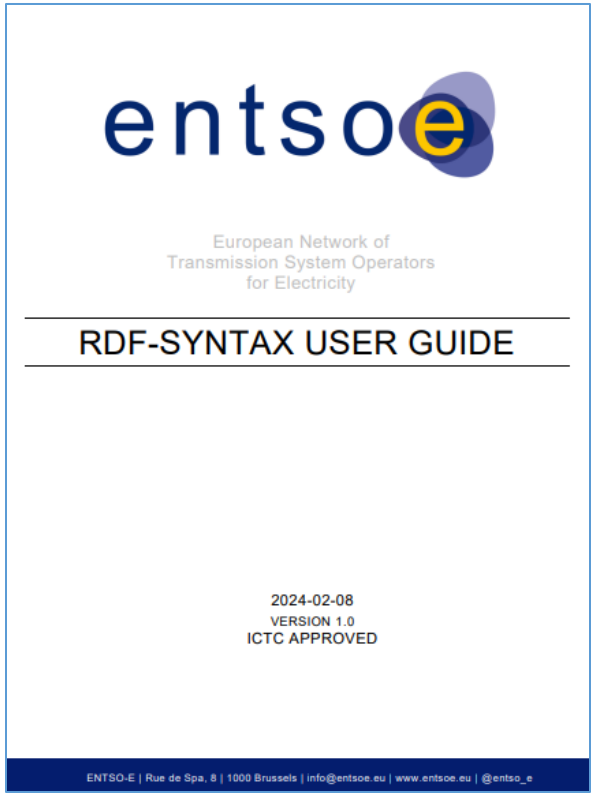


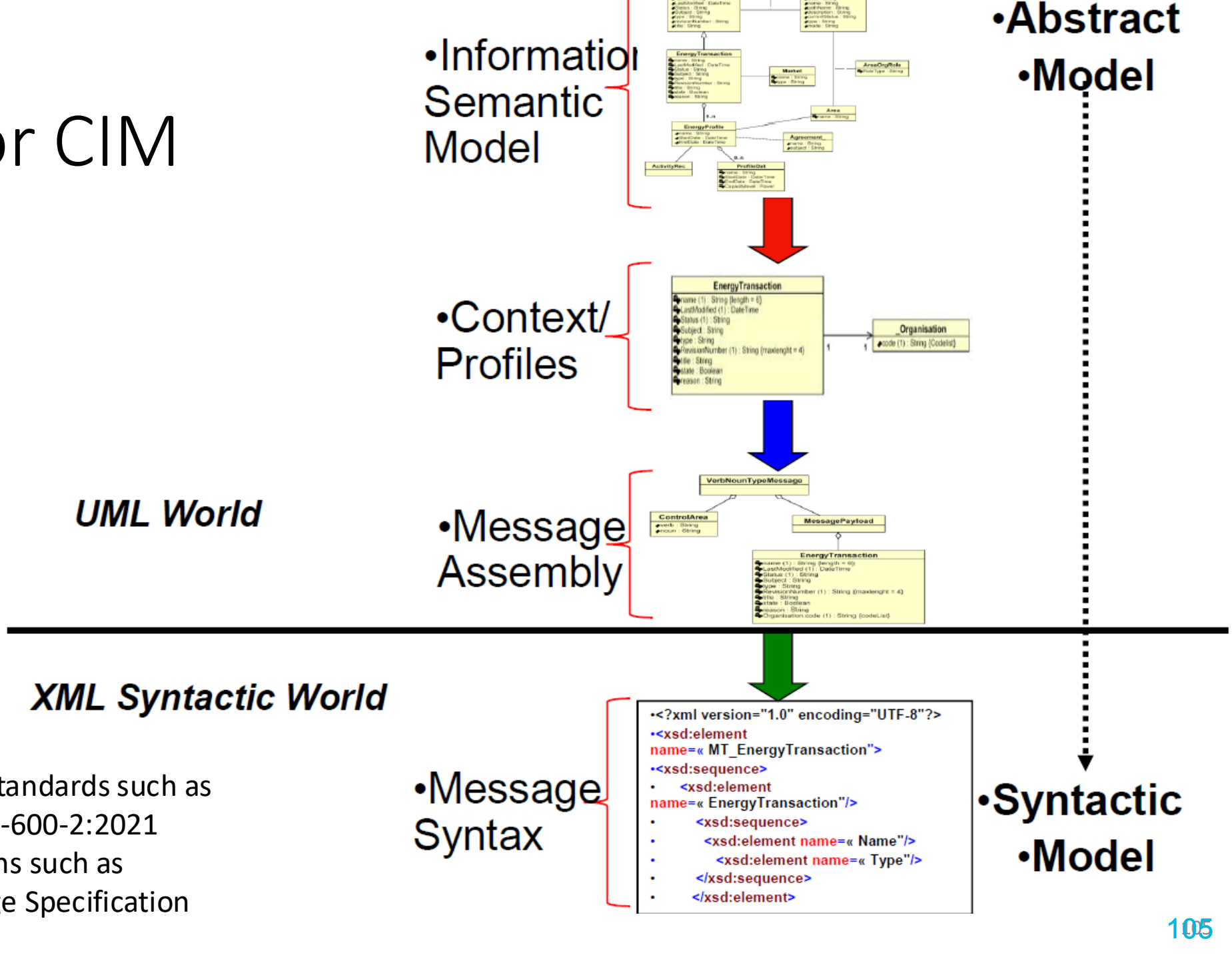
Fig. 3. Overview of a subset of classes and properties of the OEO to illustrate how they are organised inside the OEO. A black arrow denotes "is a", i.e. a subclass relation.

Booshehri, Meisam, et al. "Introducing the Open Energy Ontology: Enhancing data interpretation and interfacing in energy systems analysis." Energy and AI 5 (2021): 100074.

RDF syntax for CIM



RDF based data exchange using standards such as IEC 61970-600-1:2021, IEC 61970-600-2:2021 (CGMES) or ENTSO-E specifications such as Network Code (NC) Data Exchange Specification



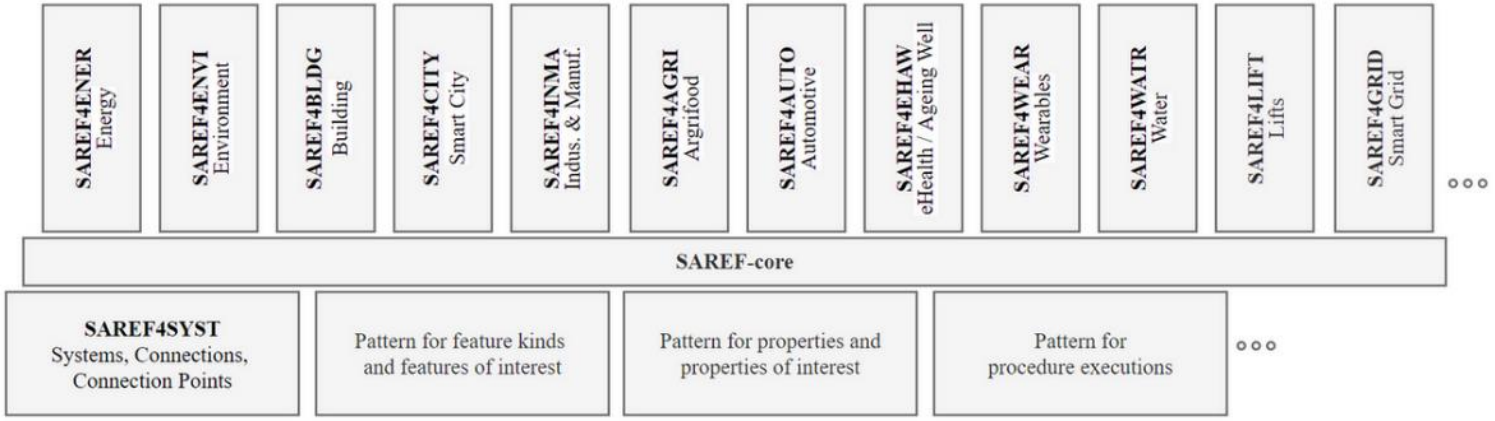
Cross-domain interoperability with the ETSI SAREF ontology

Semantic Interoperability: ETSI SAREF

- SAREF is a suite of individually versioned ontologies that contains
  - a core ontology
  - a set of reference ontology patterns that provide guidelines on how to use and extend SAREF
  - different extensions for vertical domains

  providing a mature, sustainable and **standardised framework of ontologies for IoT**

- SAREF **enables interoperability** at the semantic level **between solutions from different providers** and **among various activity sectors from the Internet of Things to Data Spaces**



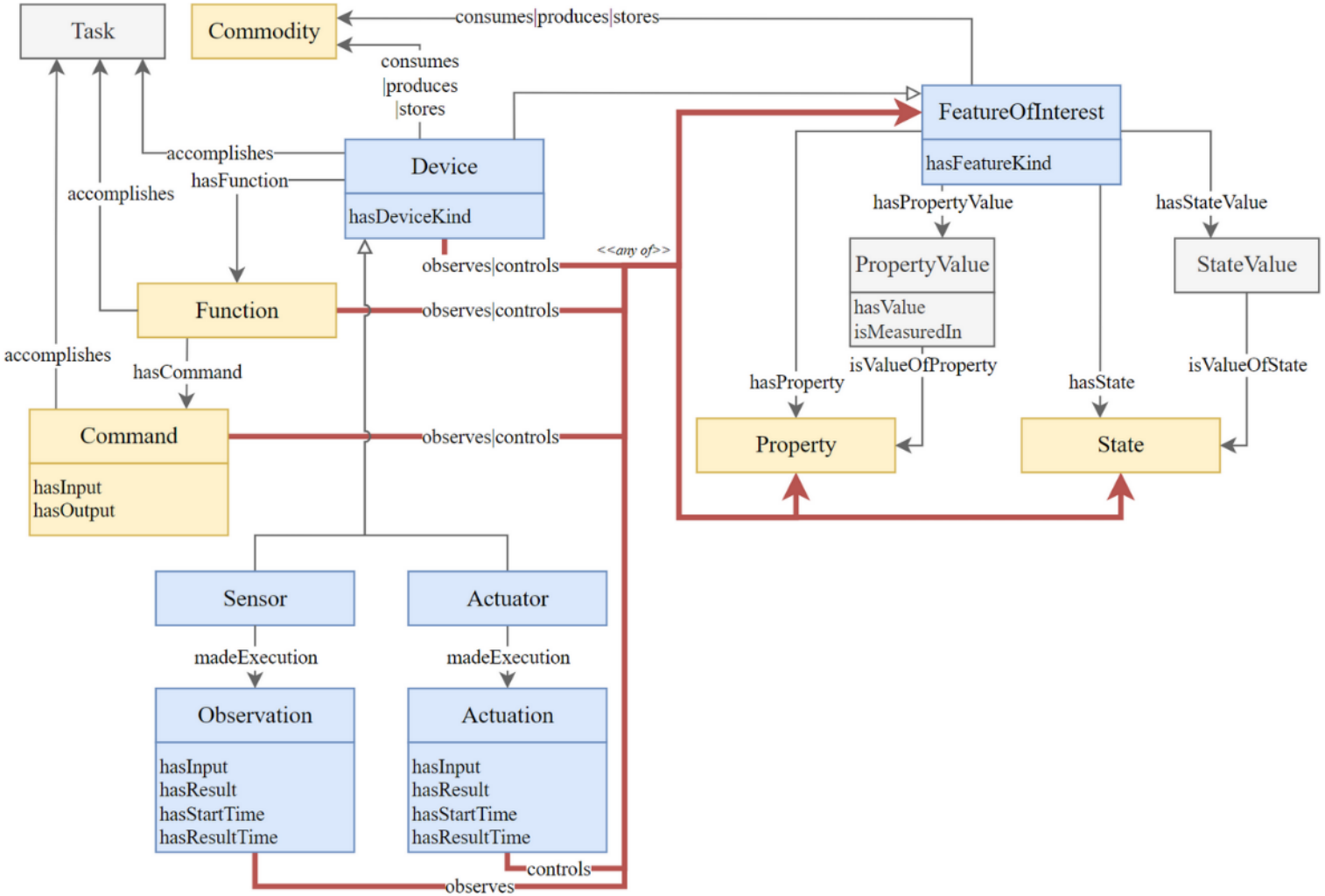
García-Castro, R., Lefrançois, M., Poveda-Villalón, M., & Daniele, L. (2023). The ETSI SAREF ontology for smart applications: a long path of development and evolution. *Energy Smart Appliances: Applications, Methodologies, and Challenges*, 183-215.

ETSI SAREF: Smart Applications REFerence Ontology

SAREF Core

SAREF in a nutshell

- Features of Interest
  - have properties and states
- Devices:
  - consume/produce/store commodities
  - accomplish tasks
  - observe properties or states of features of interest
  - control properties or states of features of interest
  - expose functions that are groups of commands
- Observations made by Sensors
- Actuations made by Actuators



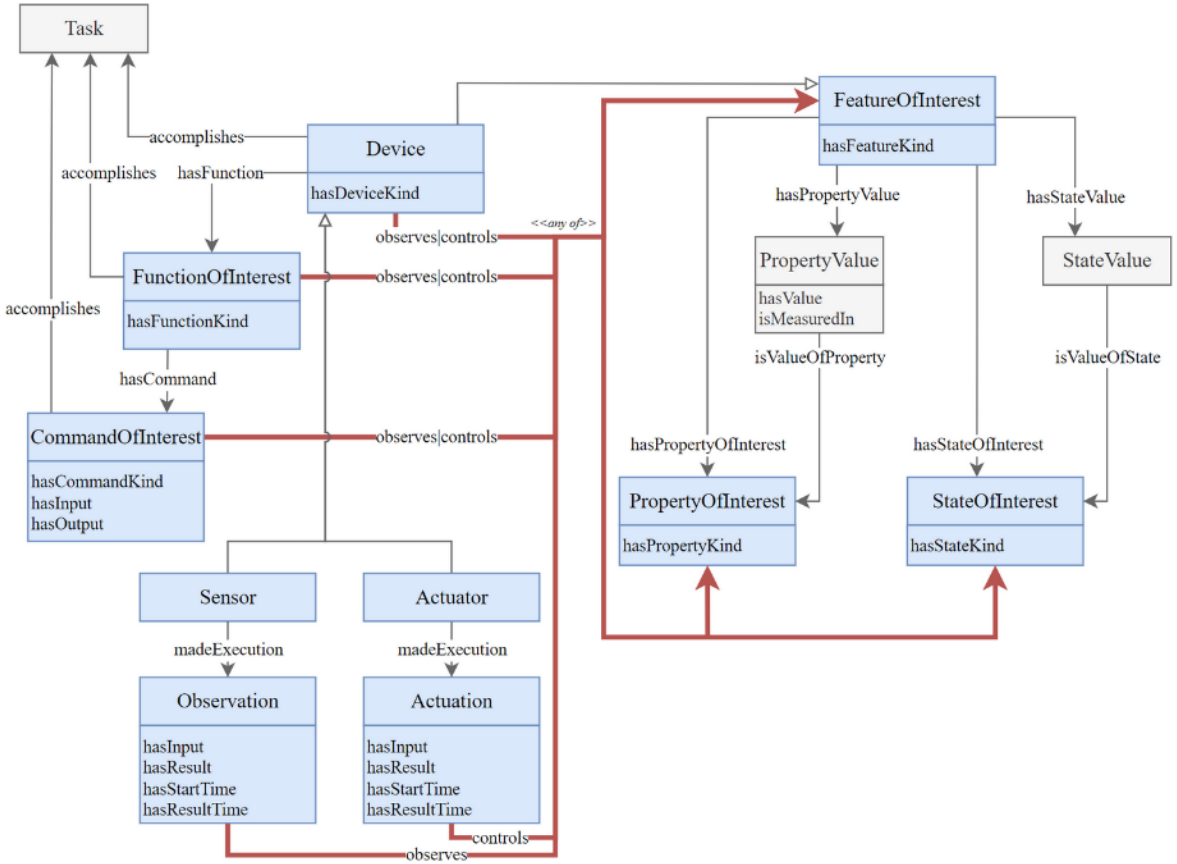
Semantic Interoperability: ETSI SAREF

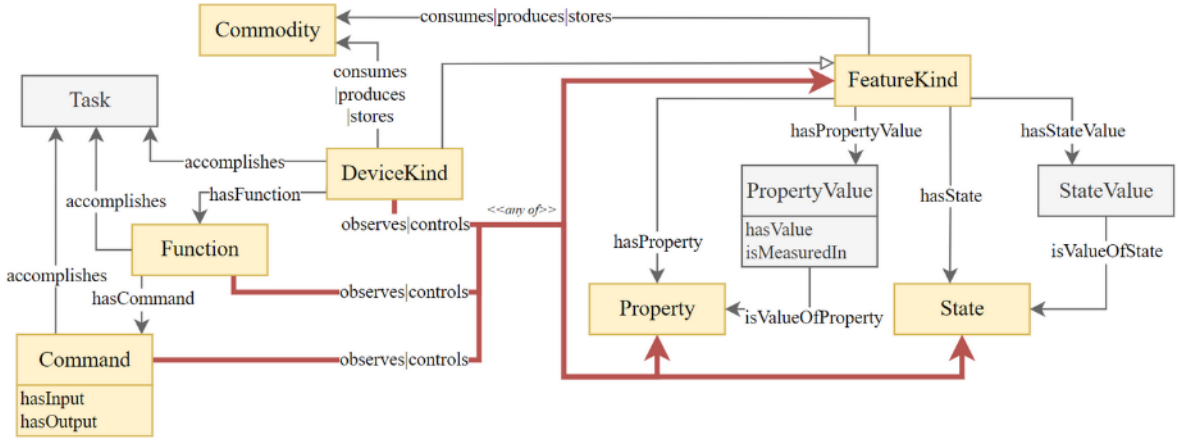
- SAREF is based on the **fundamental principles** of
  - **reuse and alignment** of concepts and relationships that are defined in **existing assets**, e.g. oneM2M base ontology, W3C® SKOS ontology, OGC® GeoSPARQL vocabulary
  - **modularity** to allow separation and recombination of different parts of the ontology depending on specific needs,
  - **extensibility** to allow further growth of the ontology,
  - **maintainability** to facilitate the process of identifying and correcting defects, accommodate new requirements, and cope with changes in (parts of) SAREF.
  - **generic versus specific entity distinction.** SAREF is designed for application developers, and also for online catalogue and taxonomy editors.

García-Castro, R., Lefrançois, M., Poveda-Villalón, M., & Daniele, L. (2023). The ETSI SAREF ontology for smart applications: a long path of development and evolution. *Energy Smart Appliances: Applications, Methodologies, and Challenges*, 183-215.

ETSI SAREF: Smart Applications REFerence Ontology

Principle of **generic versus specific entity distinction**



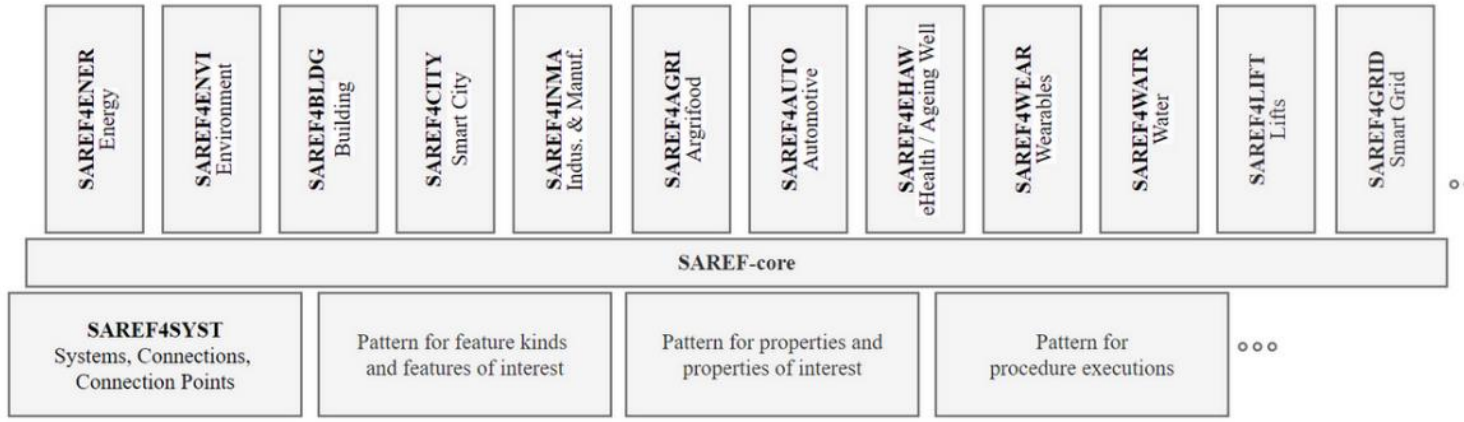


- **Smart application developers**
  - will instantiate classes such as Device and FeatureOfInterest

- **Online catalogue and taxonomy editors**
  - will populate classes such as DeviceKind, FeatureKind, Property, State, Function, Command

SAREF – Documentation and Tools

SPECIFICATIONS

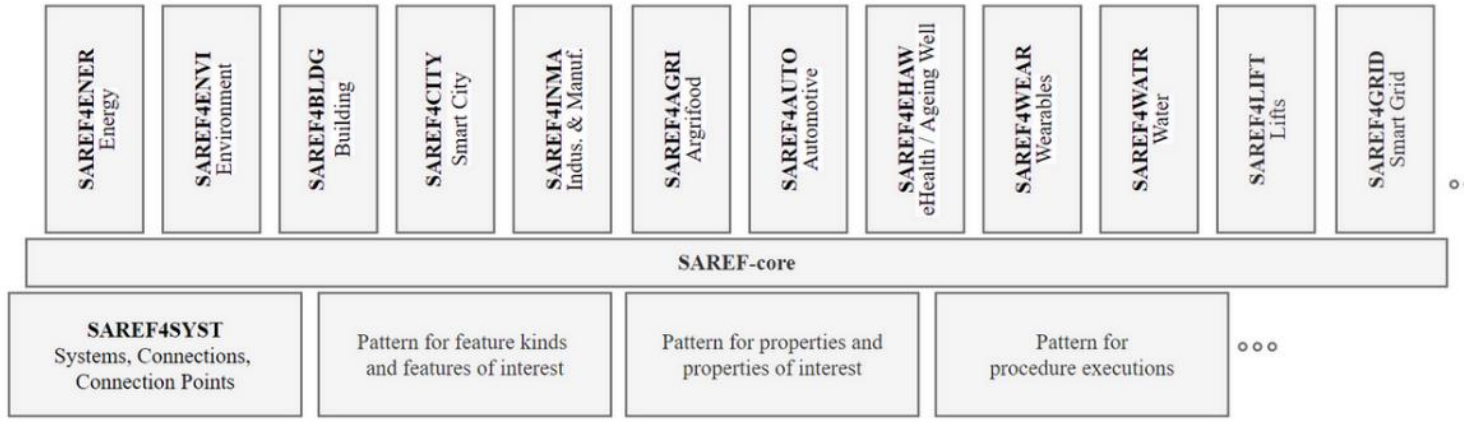


SAREF is built of a **series of Technical Specifications** specifying the SAREF suite of ontologies

EN 303 760 Guidelines for IoT Semantic Interoperability
TS 103 264 Reference Ontology and oneM2M Mapping
TS 103 548 Reference ontology patterns
TS 103 673 Development Framework and Workflow
TS 103 410-1 Energy Domain
TS 103 410-2 Environment Domain
TS 103 410-3 Building Domain
TS 103 410-4 Smart Cities Domain
TS 103 410-5 Industry and Manufacturing Domains
TS 103 410-6 Smart Agriculture and Food Chain Domain
TS 103 410-7 Automotive Domain
TS 103 410-8 eHealth/Ageing-well Domain
TS 103 410-9 Wearables Domain
TS 103 410-10 Water Domain
TS 103 410-11 Lift Domain
TS 103 410-11 Smart Grid Domain

SAREF – Documentation and Tools

SPECIFICATIONS



SmartM2M;
Smart Applications;
Reference Ontology and oneM2M Mapping

SmartM2M;
Extension to SAREF;
Part 8: eHealth/Ageing-well Domain

SmartM2M;
Extension to SAREF;
Part 1: Energy Domain

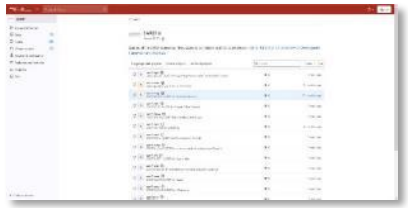
SAREF is built of a **series of Technical Specifications** specifying the SAREF suite of ontologies

TOOLS

SAREF ontology portal

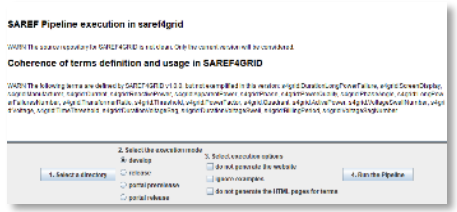
https://saref.etsi.org/

SAREF forge



https://labs.etsi.org/rep/saref/

SAREF pipeline



https://labs.etsi.org/rep/saref/saref-pipeline/

In order facilitate the SAREF suite of ontologies and to enable the SAREF community and industry stakeholders to contribute directly to the SAREF evolution **accompanying tools** have been developed in ETSI

SAREF Patterns and Dev. Framework

- ✓ ensure a homogeneous structure of the overall ontology
- ✓ speed up the development of extensions
- ✓ improve cross-domain semantic interoperability.
- ➢ SAREF ontology $\approx \sum_i f_i(x_i) + b$

patterns parameters specificities

- ➢ Different users
  - SAREF extension developer
  - On-line taxonomy/catalog editor
  - End-user application developer

explicit how-tos

- ➢ Different ways to extend
  - Sub-categorization
  - Specialization
  - Instantiation

SAREF Patterns and Dev. Framework

- ✓ ensure a homogeneous structure of the overall ontology
- ✓ speed up the development of extensions
- ✓ improve cross-domain semantic interoperability.
- ➢ SAREF ontology $\approx \sum_i f_i(x_i) + b$

patterns parameters specificities

- ➢ **Reference ontology patterns** (TS 103 548) provide guidelines on how to use and extend SAREF, to describe any kind of applications-related data/information/systems in different verticals
- ➢ **SAREF Dev. Framework** (TS 103 673) supports pattern-based development

SAREF – For the Energy domain

SAREF4ENER – SAREF for Energy Flexibility
https://saref.etsi.org/saref4ener/

Based on EN 50491-12-2

SAREF4GRID – SAREF for the Smart Grid domain
https://saref.etsi.org/saref4grid/

Based on DLMS/COSEM

ETSI SAREF: Smart Applications REFerence Ontology

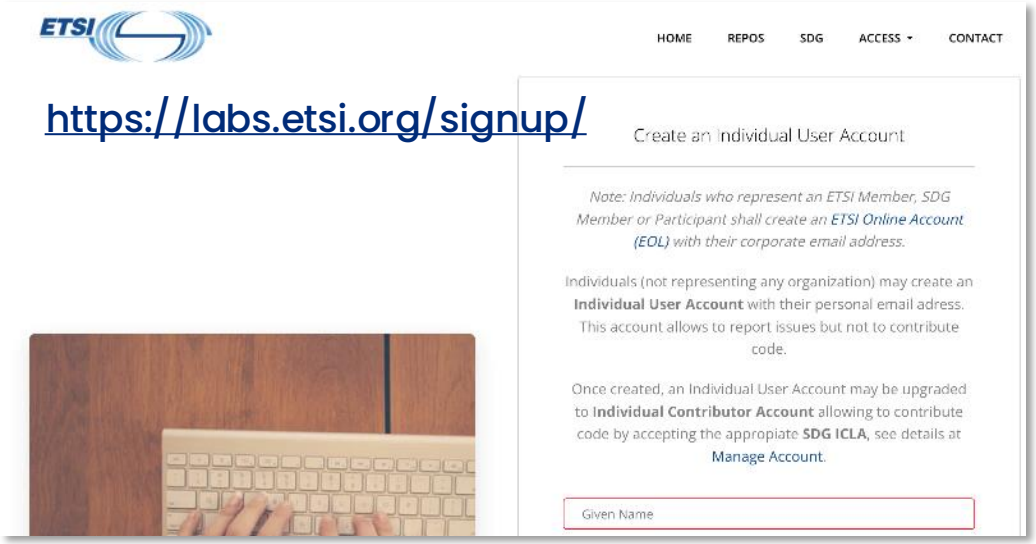
Transparency and openness of the work

- ✓ Issues, discussions, decisions publicly available on ETSI Labs

  https://labs.etsi.org/rep/groups/saref/-/issues/

- ✓ Create an Individual User Account and **weight in the discussions!**

  - **saref-portal**: homogenization and consolidation, new extension, portal…
  - **saref-core**: specific to SAREF Core
  - **saref4abcd**: specific to an extension



Drive the point home with the Data Act

European Commission

Home > Strategy > Priorities 2019-2024 > A Europe fit for the digital age > European data strategy

The European Data Act

(adopted 01/2024 – applicable 09/2025)

1)

https://ec.europa.eu/newsroom/dae/redirection/document/83517

European Commission

Home > Strategy > Priorities 2019-2024 > A Europe fit for the digital age > European data strategy

The European Data Act

(adopted 01/2024 – applicable 09/2025)

Official Journal of the European Union

EN

Series L

2023/2854

22.12.2023

REGULATION (EU) 2023/2854 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 13 December 2023

on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act)

(Text with EEA relevance)

http://data.europa.eu/eli/reg/2023/2854/oj

The Data Act will make **more data available** for use. It will set up rules on who can use and access what data for which purposes across all economic sectors in the EU.

2)

https://digital-strategy.ec.europa.eu/en/factpages/data-act-explained

European Commission

Home > Strategy > Priorities 2019-2024 > A Europe fit for the digital age > European data strategy

The European Data Act

(adopted 01/2024 – applicable 09/2025)

Chapter VIII – Interoperability

Article 33 – Essential requirements regarding interoperability of data, of data sharing mechanisms and services, as well as of common European data spaces

- Participants in data spaces that offer data or data services to other participants shall comply with the following essential requirements to facilitate the interoperability of data, of data sharing mechanisms and services, as well as of common European data spaces which are purpose- or sector-specific or cross-sectoral interoperable frameworks for common standards and practices to share:

- the dataset content, use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described, where applicable, in a machine-readable format, to allow the recipient to find, access and use the data
- the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists, where available, shall be described in a publicly available and consistent manner
- the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously, in bulk download or in real-time in a machine-readable format where that is technically feasible and does not hamper the good functioning of the connected product
- where applicable, the means to enable the interoperability of tools for automating the execution of data sharing agreements, such as smart contracts shall be provided

COMMISSION IMPLEMENTING DECISION on a standardisation request to the European standardisation organisations as regards a European Trusted Data Framework in support of Regulation (EU) 2023/2854 of the European Parliament and of the Council

Table 1: List of new European standards and European standardisation deliverables to be drafted and deadlines for their adoption

	Reference information	Deadline for the adoption by the ESOs
1.	Harmonised standards on Trusted Data Transactions Part 1: Terminology, concepts and mechanisms	1 June 2026
2.	Harmonised standards on Trusted Data Transactions Part 2: Trustworthiness requirements	1 November 2026
3.	Harmonised standards on Trusted Data Transactions Part 3: Interoperability requirements	1 May 2027
4.	Technical specification(s) on a data catalogue implementation framework	1 March 2026
5.	Technical specification(s) on an implementation framework for semantic assets	1 September 2026
6.	European standard on a quality framework for internal data governance	1 March 2027
7.	Technical specification(s) on a maturity model for Common European Data Spaces	1 September 2026

COMMISSION IMPLEMENTING DECISION on a standardisation request to the European standardisation organisations as regards a European Trusted Data Framework in support of Regulation (EU) 2023/2854 of the European Parliament and of the Council

Table 1: List of new European standards and European standardisation deliverables to be drafted and deadlines for their adoption

	Reference information	Deadline for the adoption by the ESOs
1.	Harmonised standards on Trusted Data Transactions Part 1: Terminology, concepts and mechanisms	1 June 2026
2.	Harmonised standards on Trusted Data Transactions Part 2: Trustworthiness requirements	1 November 2026
3.	Harmonised standards on Trusted Data Transactions Part 3: Interoperability requirements	1 May 2027
4.	Technical specification(s) on a data catalogue implementation framework	1 March 2026
5.	Technical specification(s) on an implementation framework for semantic assets	1 September 2026
6.	European standard on a quality framework for internal data governance	1 March 2027
7.	Technical specification(s) on a maturity model for Common European Data Spaces	1 September 2026

Data catalogue implementation framework

This (these) technical specification(s), entry 4 in table 1 of Annex I, shall provide a framework for standardised catalogue metadata, in support of the findability of data within and across data spaces.

The development of the technical specification(s) shall take into account the Interoperable Europe solutions based on the W3C Data Catalogue Vocabulary (DCAT) standard, in particular the DCAT-AP, DCAT-AP-HVD and GeoDCAT-AP profiles[5]

The technical specification(s) shall:

(a) set out the common catalogue metadata, to be applied across all common European data spaces;

(b) establish rules on the setting out of domain-specific catalogue metadata, to be applied in selected common European data spaces;

EUROPEAN COMMISSION

COMMISSION IMPLEMENTING DECISION on a standardisation request to the European standardisation organisations as regards a European Trusted Data Framework in support of Regulation (EU) 2023/2854 of the European Parliament and of the Council

Table 1: List of new European standards and European standardisation deliverables to be drafted and deadlines for their adoption

	Reference information	Deadline for the adoption by the ESOs
1.	Harmonised standards on Trusted Data Transactions Part 1: Terminology, concepts and mechanisms	1 June 2026
2.	Harmonised standards on Trusted Data Transactions Part 2: Trustworthiness requirements	1 November 2026
3.	Harmonised standards on Trusted Data Transactions Part 3: Interoperability requirements	1 May 2027
4.	Technical specification(s) on a data catalogue implementation framework	1 March 2026
5.	Technical specification(s) on an implementation framework for semantic assets	1 September 2026
6.	European standard on a quality framework for internal data governance	1 March 2027
7.	Technical specification(s) on a maturity model for Common European Data Spaces	1 September 2026

Semantic assets implementation framework

This (these) technical specification(s), entry 5 in table 1 of Annex I, shall provide a framework for common, open vocabularies, classification schemes, taxonomies, code lists and ontologies, in support of the interpretation and analysis of shared data within and across data spaces.

Existing standards-based frameworks, in particular the Core Vocabularies and the Asset Description Metadata Schema Application Profile/ADMS-AP (EC - SEMIC), the Asset Administration Shell (IEC), the ISO/IEC 19115 metadata standards, the European Commission *countries and territories* reference data asset[6] for geospatial data, and the Smart Applications REFerence/SAREF Ontology (ETSI), shall be taken into account.

The technical specification(s) shall:

(a) specify criteria for the selection of semantic assets;

(b) specify methods for the semantic annotation of shared data, the detailed metadata, based on the semantic assets mentioned under point (a).

EUROPEAN COMMISSION

Maxime Lefrançois

Associate Professor
IMT - MINES Saint-Étienne - Institut Henri Fayol
Member of Laboratoire LIMOS - UMR CNRS 6158
https://maxime-lefrancois.info