FEMS 2025 EUROMAT

18th European Congress and Exhibition on Advanced Materials and Processes

Granada,
14 – 18 September 2025

EMMO Ontology: enabling AI-based innovative advanced materials development: the CoBRAIN Knowledge Base for Hardmetal Thermal Spraying Coatings

Gerhard Goldbeck
EMMC ASBL

Emanuele Ghedini,
University of Bologna

FEMS 2025 EUROMAT

Overview

1. Introduction to EMMC
2. Introduction to the CoBrain project and the challenge to be addressed
3. State of the art in data documentation and management and materials domain requirements
4. Conceptualisation: expressing data and knowledge via ontologies
5. The CoBrain experience
6. Lessons learnt and outlook
7. Acknowledgements

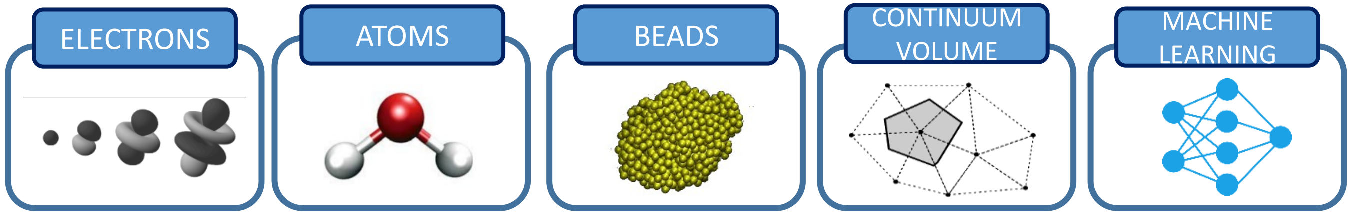
EMMC

aid4 GREENEST

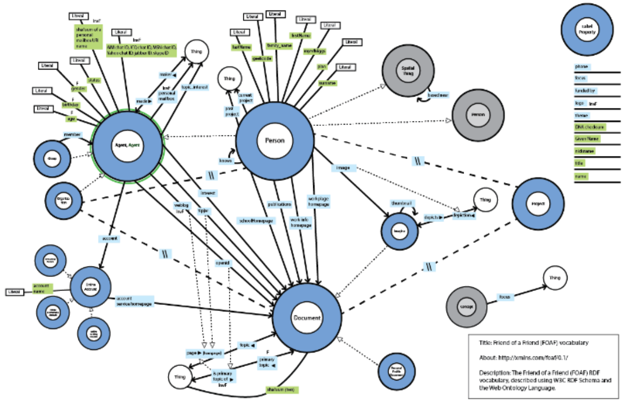
CoBRAIN

EMMC: modelling the materials world

- **ALL types of materials models**: physics and data based, incl AI



- **ALL** chemical/material and application fields



- **Data models and ontologies**, shared conceptualisations based on materials science foundations

Now including Materials Characterisation

- **Harmonisation and standardisation.**

EMMC working with the community

CEN WORKSHOP AGREEMENT

CWA 17815

January 2025

ICS 01.040.07; 07.120; 17.020

Supersedes CWA 17815:2021

English version

Materials characterization - Terminology and structured documentation

Terminologies and metadata

EMMC 2025

Accelerated Innovation and Sustainability by a Materials Modelling and Data Ecosystem

8-10 April 2025, TU Wien, Vienna (Austria)

International Workshop

https://emmc.eu/emmc-2025/

Training Resources

Home | Training Resources

The Value of Materials Modelling

A collection of real-world challenges and their solution

https://emmc.eu/simulation-success-stories/

EMMO Ontology

https://emmc.eu/emmo/

Github for collaborations

European Material Modelling Council

Repositories for collaboration on codes beteen EMMC-related projects

5 followers https://emmc.eu @EMMC_ASBL contact@emmc.eu

https://github.com/EMMC-ASBL/

zenodo

EMMC ASBL

https://zenodo.org/communities/emmc

EMMC Services for projects

https://emmc.eu/services/

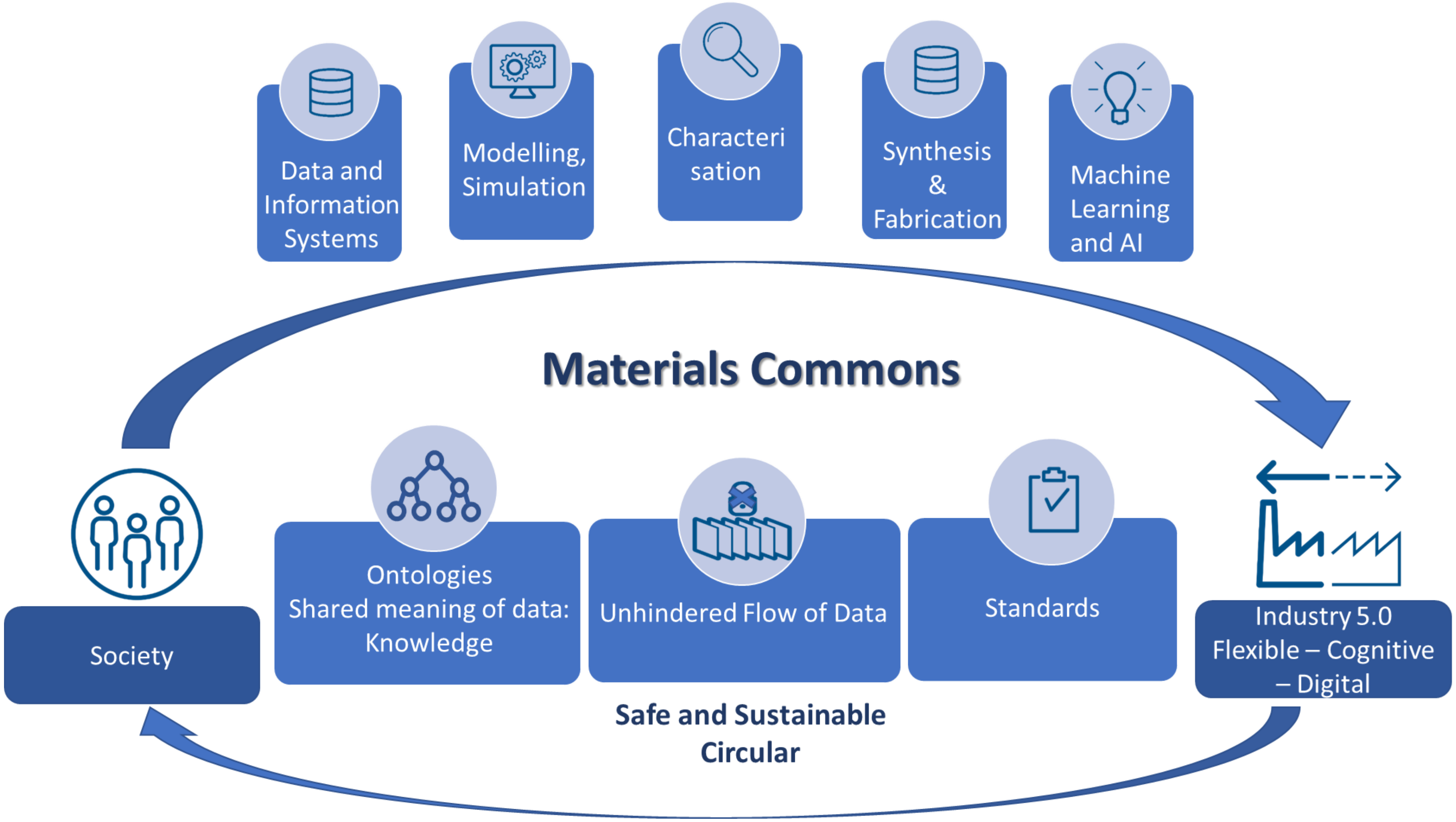
EMMC as services **contractor** supports

- Collaboration, clustering via workshops, surveys, reports
- Uptake of FAIR data standards and use of ontologies

Please make sure to include EMMC services in your project proposal

iENTRANCE MatCHMaker Materials Characterisation & Modelling MaMMoS DigiCell aid4 GREENEST

Virtuous cycle of knowledge generation and exploitation



Materials Commons

- An **interconnected resource for all.**

- **Facilitating**
  - workflow integration
  - data documentation
  - semantic interoperability

  - …and AI

Midsummer Commons, Cambridge, with cows and cyclist

INTEGRATED COMPUTATIONAL/EXPERIMENTAL MATERIAL ENGINEERING OF THERMAL SPRAY COATINGS

Co-BRAIN

- **Protective coatings** against wear and corrosion
- Current technologies: drawbacks in terms of sustainability
  - ✓ Electroplated Cr is a safe material, but its deposition involves the use of **carcinogenic Cr^{6+}** compounds subject to authorisation under REACH Annex XIV
  - ✓ Electroplated or electroless Ni(P) or Ni(B) based layers also use hazardous raw materials and result in coatings that might be less safe
  - ✓ Thermal spray WC-Co-based coatings utilize carcinogenic materials – Co – as well as critical raw materials – Co, W

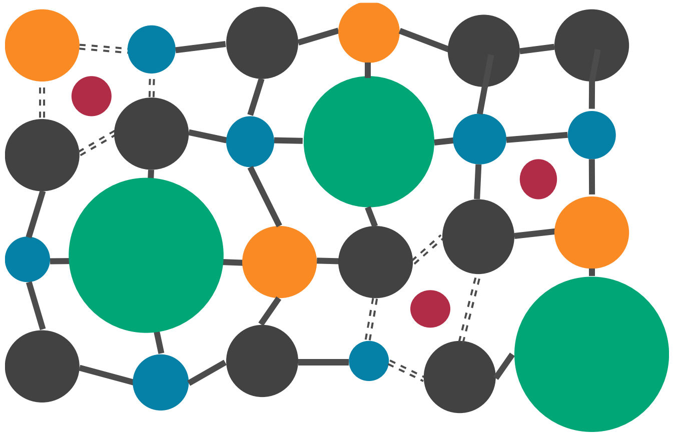
⇒ **Identify novel formulations**, use **thermal spraying (**versatility and low environmental impact) **to produce alloy and hardmetal coatings free of toxic and critical materials.**

Funded by the European Union

Funded by the European Union under the GA no 101092211. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Topic

High Entropy Alloys - HEA



Materials space

Multi-principal element metal matrix and carbides

Elements for HEA metal binder

Al	Si	Ti	V	Cu
Cr	Mn	Fe	Ni	Co

Elements for hard phase

Ti C

B	Nb	Mo
Hf	Ta	W

126 equiatomic combinations of 4 or 5 elements

62370 cermets (without composition optimization)

- **High-entropy effect**

Multi-element systems often consist of **one (HEAs) or two main phases**, instead of a wider range of binary or ternary solid solutions and intermetallics, because of the stabilising entropy of a multi-element random solid solution

($\Delta G = \Delta H - T \cdot \Delta S$)

- **Lattice strain effect**

Increased strength: mismatched atomic radii cause lattice distortions.

- **Sluggish diffusion effect**

Better high-temperature stability: The distorted lattice hinders long-range diffusion

- **"Cocktail" effect**

Non-linearities and unexpected synergistic effects may sometimes yield **somewhat unpredictable outcomes**

→ Need to combine experimental development, physical modelling & artificial intelligence

CoBRAIN Partners

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

access

AeonX

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

BALance® Technology Consulting

exelisis

MBN nanomaterialia®

obz

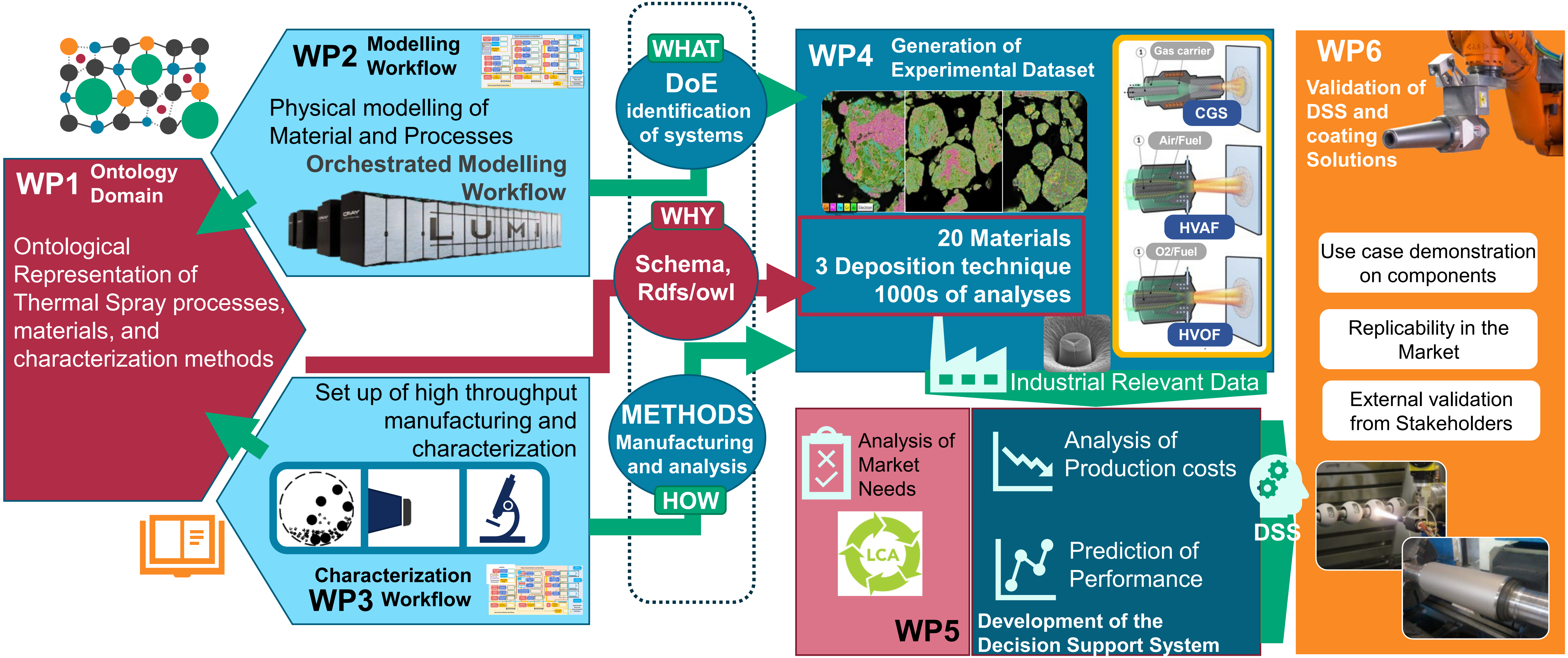
TITOMIC

UNIVERSITAT DE BARCELONA

ROMA TRE UNIVERSITÀ DEGLI STUDI

VTT

The Project

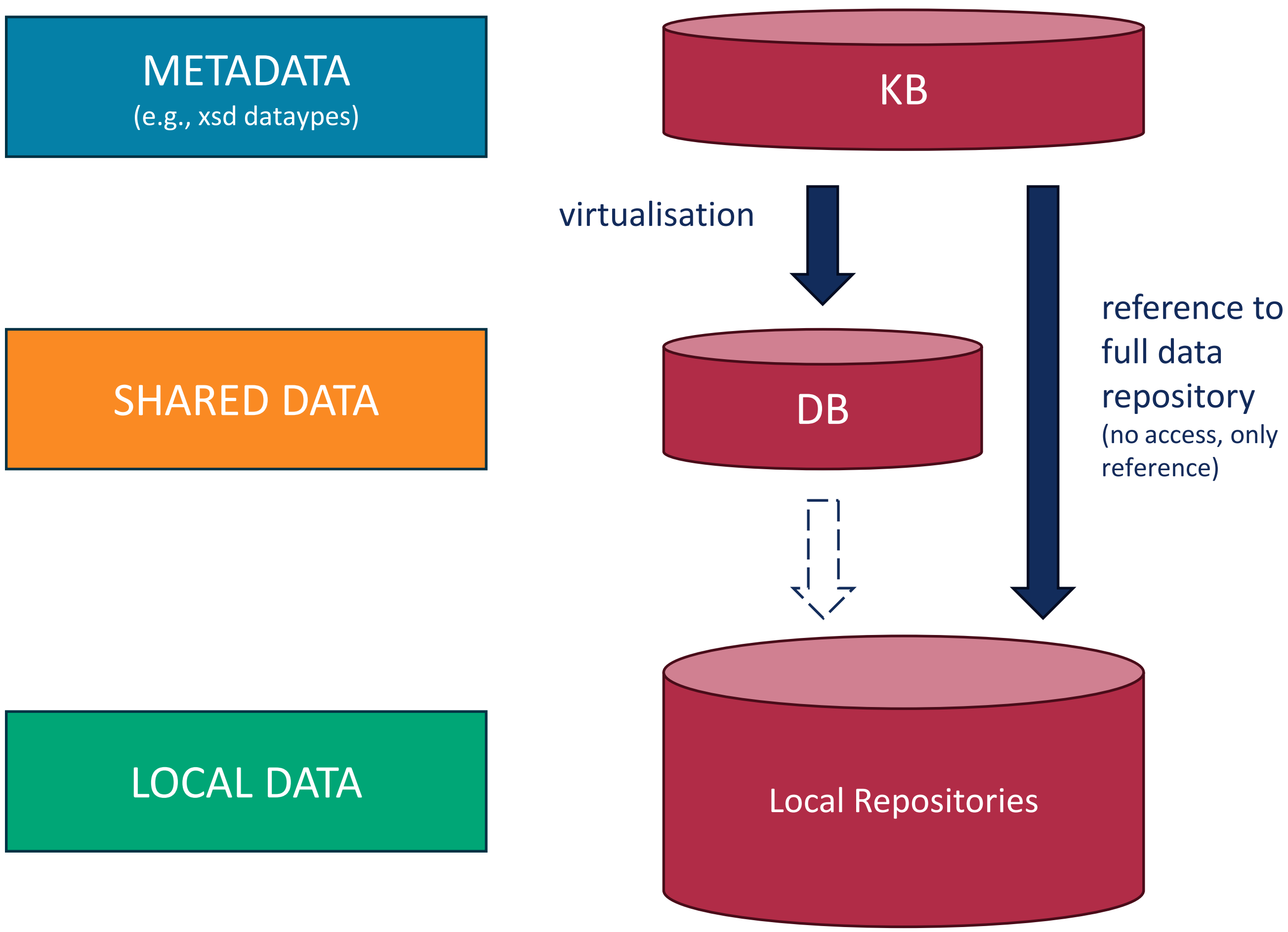


WP1 Ontology Domain
Ontological Representation of Thermal Spray processes, materials, and characterization methods
WP2 Modelling Workflow
Physical modelling of Material and Processes
Orchestrated Modelling Workflow
LUMI
Set up of high throughput manufacturing and characterization
Characterization WP3 Workflow
WHAT
DoE identification of systems
WHY
Schema, Rdfs/owl
METHODS
Manufacturing and analysis
HOW
WP4 Generation of Experimental Dataset
20 Materials
3 Deposition technique
1000s of analyses
Gas carrier
CGS
Air/Fuel
HVAF
O2/Fuel
HVOF
Industrial Relevant Data
Analysis of Market Needs
LCA
WP5
Analysis of Production costs
Prediction of Performance
Development of the Decision Support System
DSS
WP6
Validation of DSS and coating Solutions
Use case demonstration on components
Replicability in the Market
External validation from Stakeholders

Intended knowledge base architecture

Co-BRAIN

Original Architecture



SHARED DATA and METADATA

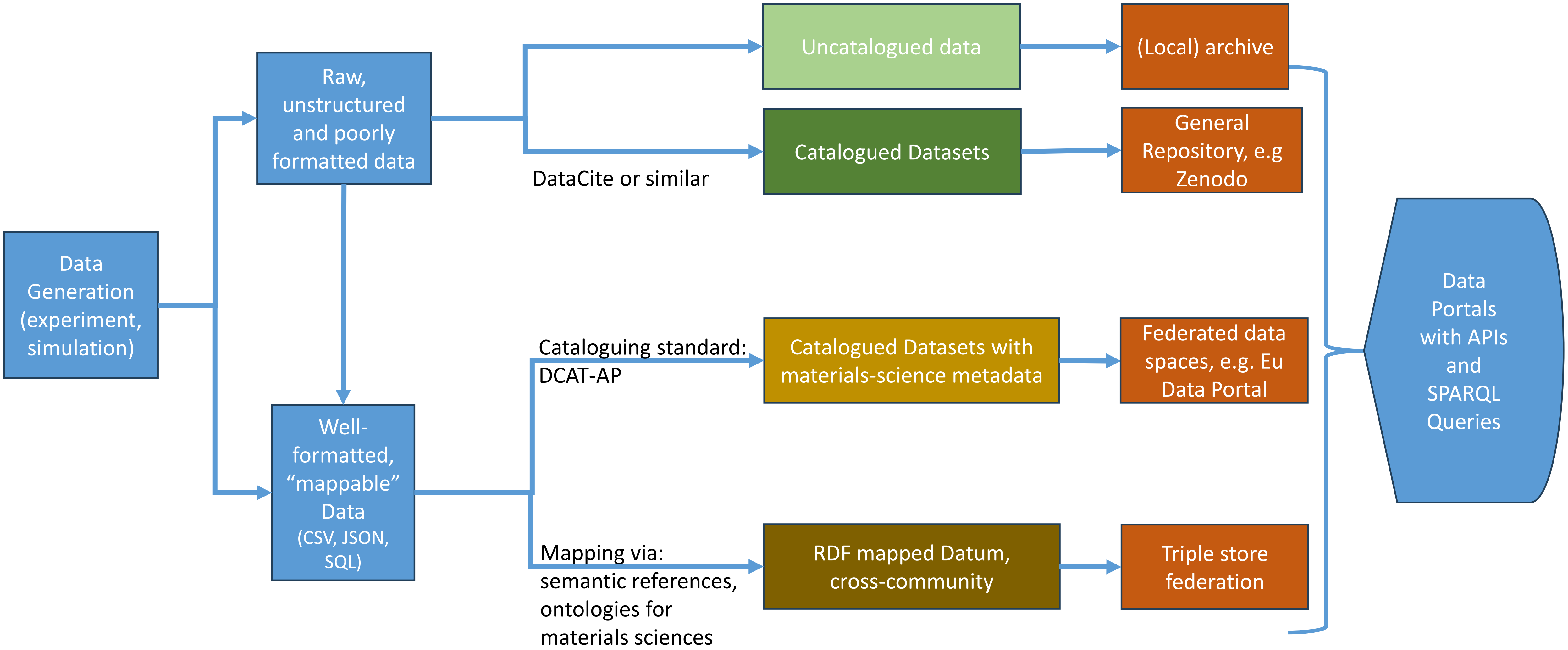
Field	Description
Sample ID	String: Sample name and identification number of test.
Sample description	String: sample composition and production information.
Test definition	String: Type of test and instrument name (i.e. Nanoindentation-Pillar splitting, MTS G200).
SOP file name	String: indication of the file name containing parameters of measurement (i.e. acquisition rate, max load, loading rate, approach speed etc.) .
Measurement Date Time	dd/mm/aaaa hh:mm:ss
Indenters	String: type of the indenter (i.e. Berkovich); serial number; material of the indenter.
Curves	Indentation curve: Load applied [N] vs penetration depth [m].
P_C	**mN,** the value of the critical load, identified from Indentation curve: (Load applied [N] vs penetration depth).
K_{IC}	**MPa√m,** the value of the fracture toughness calculated from the value of the critical load.

LOCAL DATA

Field	Description
Sample ID	String: Sample name and identification number of test.
Measurement file	**.mss**, located in UNIROMA3's servers, contains all the information related to the measurement and the complete results.
Indentation Data	**.xls**, located in UNIROMA3's servers, contains data, exported from .mss, needed for fracture toughness calculation (i.e. applied load [mN] vs penetration depth [nm]).

Data integration architecture

Inspired by CONNECT-NM
https://www.connect-nm.eu/



Background: semantic web technologies

- **A flexible, interconnected data model. Data linked as triples:**
  - **Subject-predicate (property)-object,**
  - **e.g. EMMC (an organisation) “is registered in” Brussels (a city)**
- **Saves time, cost, and improves maintainability**
  - No fixed schema
  - Open World Assumption (OWA):

  new facts can be easily included.

https://devopedia.org/semantic-web

The path towards FAIR

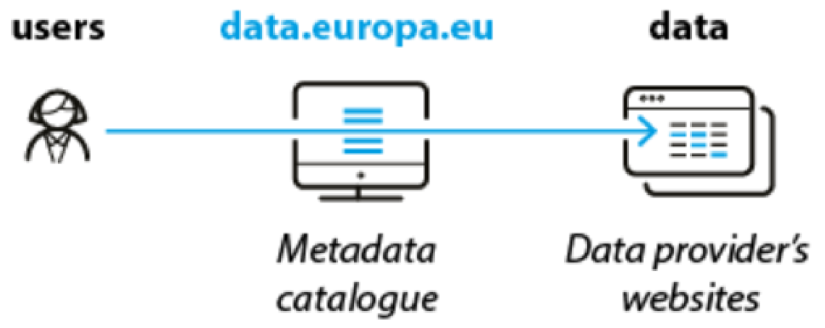
- Data management: structured data, databases
- Data cataloguing
  - Findable, Accessible
- Data exchange protocols: APIs, SPARQL, etc
  - Accessible, Reusable
- Data models and ontologies: semantic specifications
  - Interoperable

Datasets Cataloguing

- DCAT: W3C standard metadata for datasets
- DCAT-AP (Application Profiles):
  - common specification for public sector datasets
  - constraints and usages on DCAT properties and classes
  - additional properties, usage of controlled vocabularies
  - Extensions to other domains, e.g. GeoDCAT-AP: spatio-temporal
  - Enables building interconnected data spaces

Reference:
An introduction to DCAT-AP and LDES at EU Open Data Days, March 2025
https://data.europa.eu/sites/default/files/report/Georges%20Lobo%20%26%20Pavlina%20Fragkou.

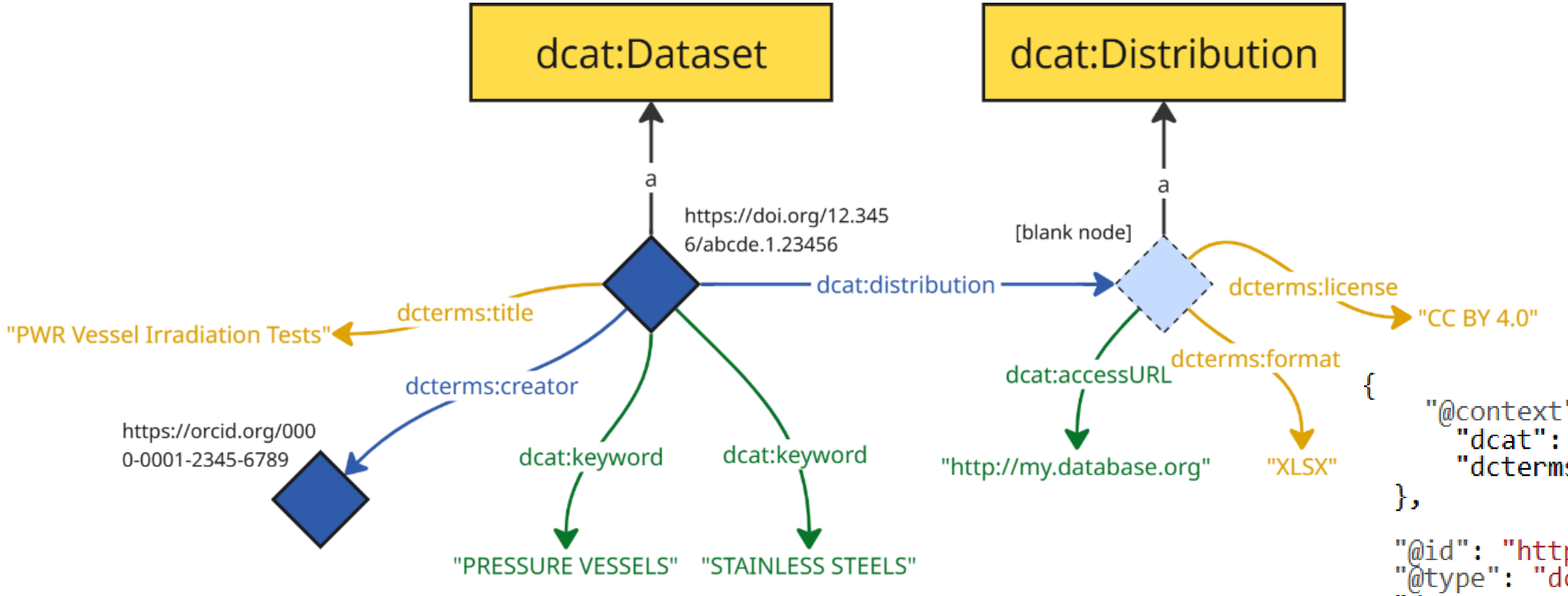
Data.europa.eu is a **catalogue of metadata** providing a single point of access to open data from European countries and EU institutions' for reuse.



Anyone can easily search, explore, link, download and reuse the data for commercial or non-commercial purposes.

https://data.europa.eu/en/publications/datastories/linking-data-dataeuropaeu

Example: Nuclear engineering test data



JSON-LD serialisation


```
{
  "@context": {
    "dcat": "http://www.w3.org/ns/dcat#",
    "dcterms": "http://purl.org/dc/terms/"
  },

  "@id": "https://doi.org/12.3456/abcde.1.23456",
  "@type": "dcat:Dataset",
  "dcterms:creator": {
    "@id": "https://orcid.org/0000-0001-2345-6789"
  },
  "dcat:keywords": ["PRESSURE VESSELS", "STAINLES STEELS"],
  "dcterms:title": "PWR Vessel Irradiation Tests",
  "dcat:distribution": {
    "@type": "dcat:Distribution",
    "dcat:accessURL": "http://my.database.org",
    "dcterms:license": "CC BY 4.0",
    "dcterms:format": "XLSX"
  }
}
```


Only "application" metadata are Keyword and Title

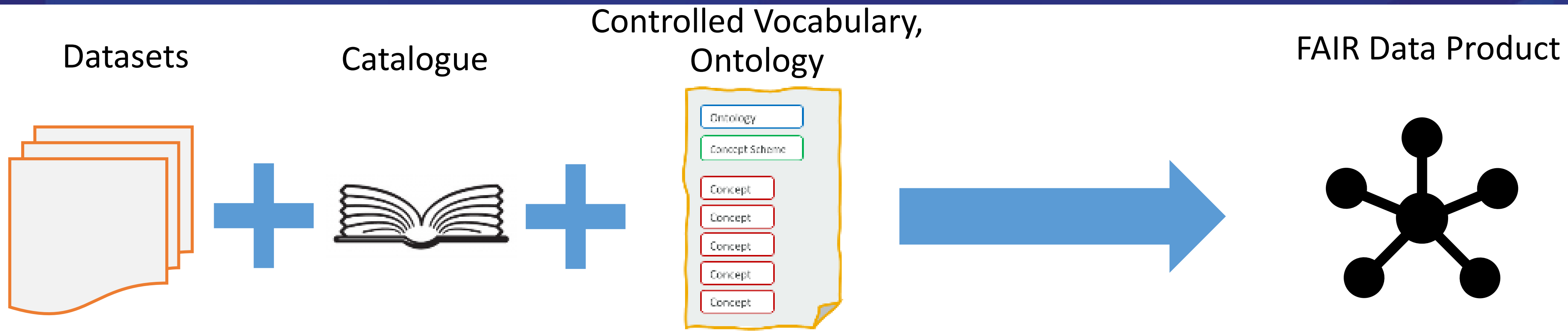
Need domain metadata and controlled vocabularies for materials :
a Materials DCAT-AP !

Knowledge Organisation Challenges in Materials Sciences

- Lack of agreed conceptualisation and terminologies
  - no equivalent to IUPAC Goldbook
  - basic terms (e.g. material vs chemical substance) are not standardised.
- Lack of coverage of materials in widely used schema (Schema.org, DCAT-AP)
- Poor and inconsistent coverage of materials terms in ontologies*
- Reflects the EOSC Semantic Interoperability Challenges:
  - *Lack of common explicit definitions about the terms that are used by user communities*
  - *Lacking common semantic artefacts across sub-domains*

*** Review and Alignment of Domain-Level Ontologies for Materials Science**
Anne De Baas, Pierluigi Del Nostro, Jesper Friis, Emanuele Ghedini, Gerhard Goldbeck, Ilaria Maria Paponetti, Andrea Pozzi, Lan Yang, Arkopaul Sarkar, Francesco Antonio Zaccarini and Daniele Toti (2023) https://ieeexplore.ieee.org/document/10296887

Minimum viable FAIR data
Credit: Ben Gardner, AstraZeneca



- **Findable**: data(sets) should be registered and discoverable in a Data Catalogue
- **Accessible**: mechanism in place for requesting and receiving data
- **Interoperable**: data mapped to agreed terms (Controlled Vocab) and instances have PIDs
- **Reusable**: Documentation describing the data (and the entities they related to): ontology

Data and Controlled Vocabularies
Putting Interoperability into FAIR

"Dirty data"

Study	Indication	Drug
D1234C00001	Non small cell lung cancer	Tagrisso
ADORA	NSCLC	Osimertinib
CP11278-CMA33G	Diabetes type 2	Forxiga

- Inconsistent identifiers & terms
- Column values can be concatenated
- etc

Data Transform

Term

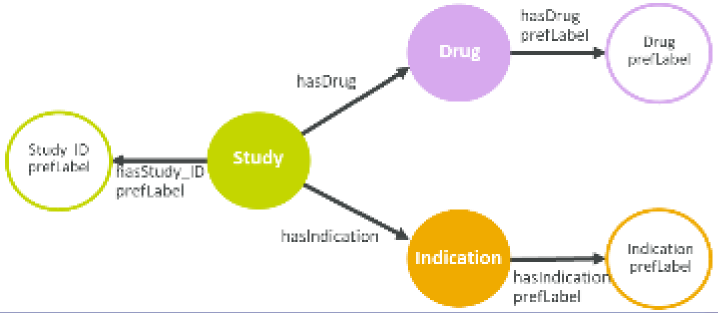
prefLabel & PID

Controlled Vocabularies

Interoperable data

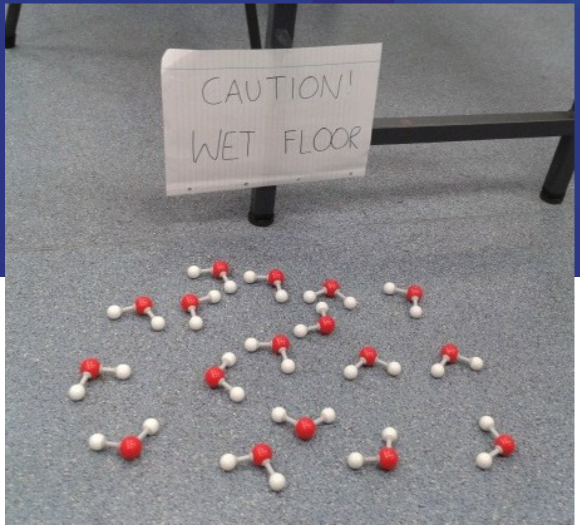
Study_ID	Study_ID__URI	Indication	Indication__URI	Drug	Drug__URI
D1234C00001	https://pid.astrazeneca.com/1/12345	Non small cell lung cancer	https://pid.astrazeneca.com/Indication/23456	Tagrisso	https://pid.astrazeneca.com/Product/965723
D1234C00012	https://pid.astrazeneca.com/1/48373	Non small cell lung cancer	https://pid.astrazeneca.com/Indication/23456	Tagrisso	https://pid.astrazeneca.com/Product/965723
D4568L00007	https://pid.astrazeneca.com/1/97538	Diabetes type 2	https://pid.astrazeneca.com/Indication/9857	Forxiga	https://pid.astrazeneca.com/Product/853584

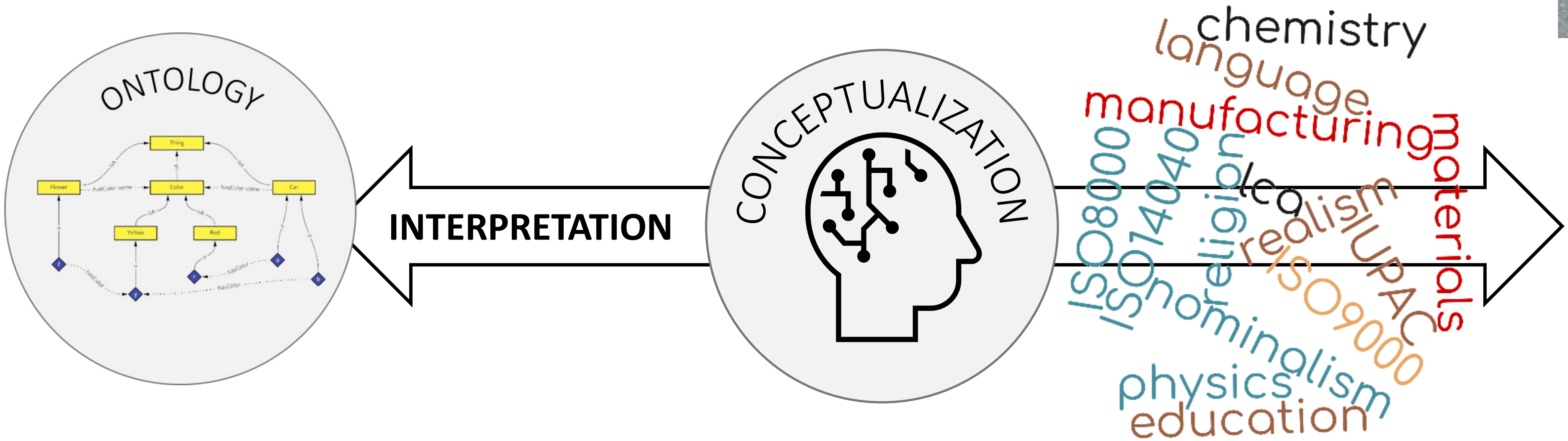
- Use of PIDs and URIs
- PrefLabels and Controlled Vocabulary
- Use common files format CSV, JSON, etc
- Graph enabling (rdf triples)



Credit: Ben Gardner, AstraZeneca

Data are more than “numbers”



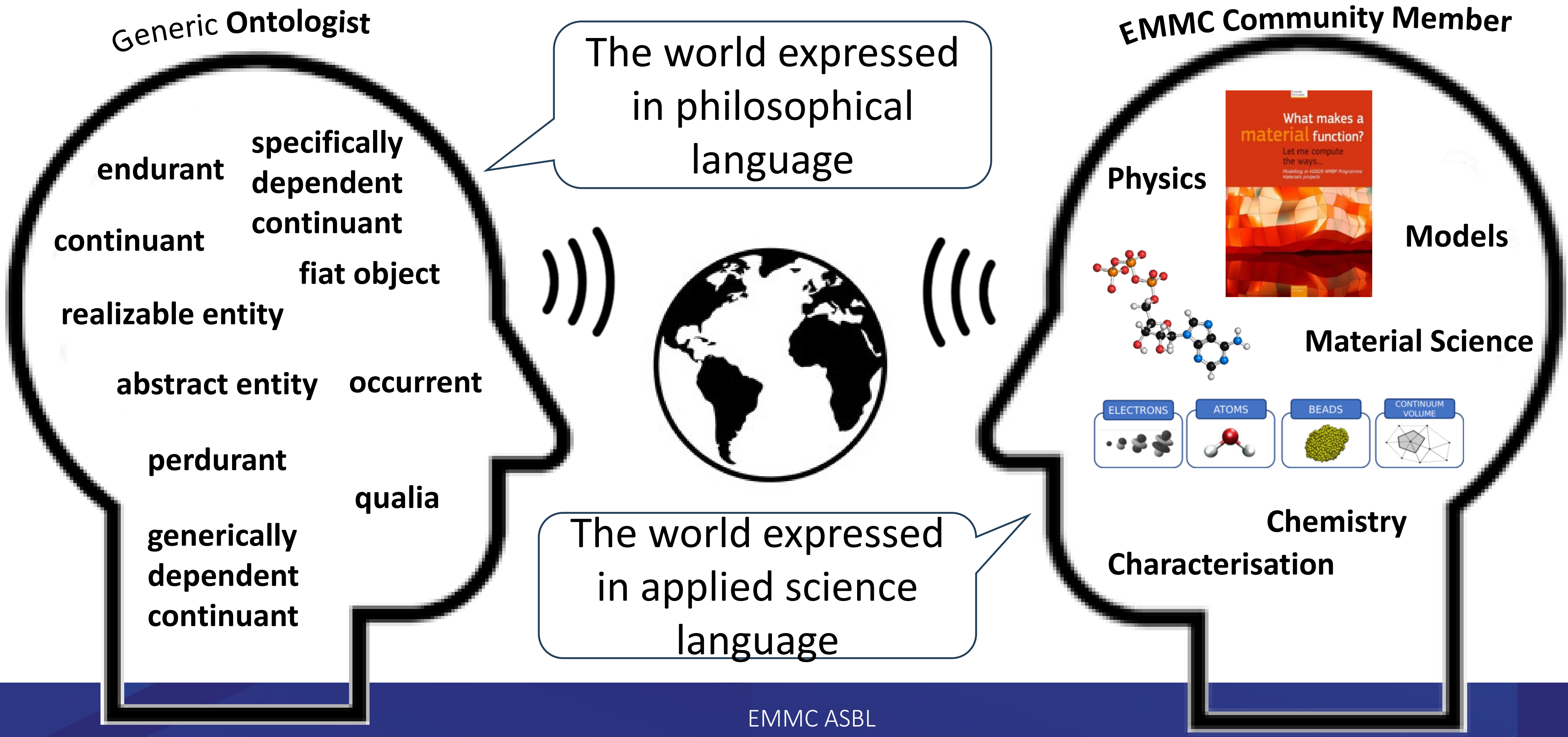


Several conceptualizations for **the same things**

Different things can have the same label

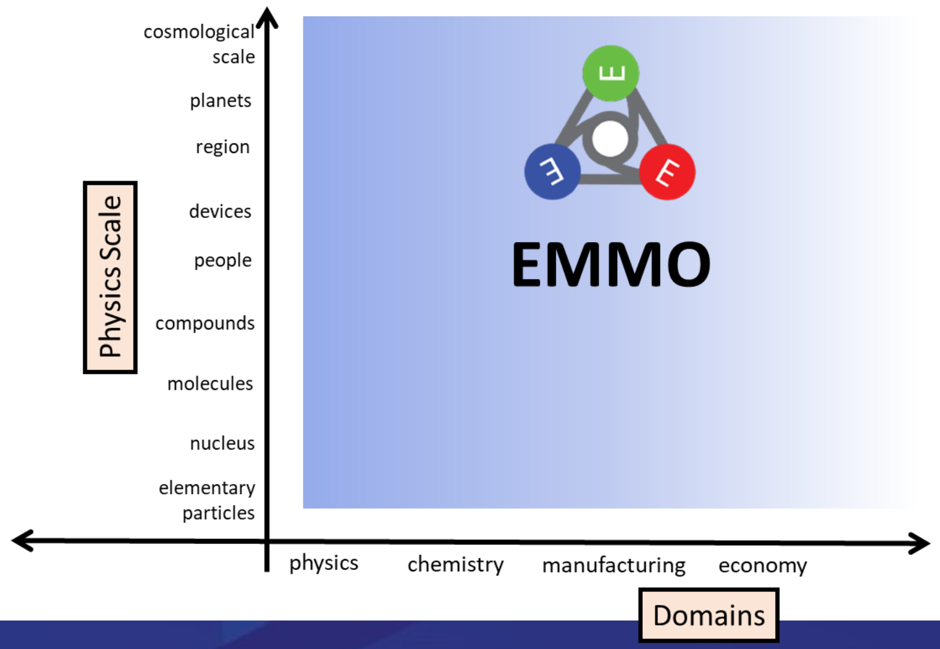
EMMO: a foundational ontology for materials sciences

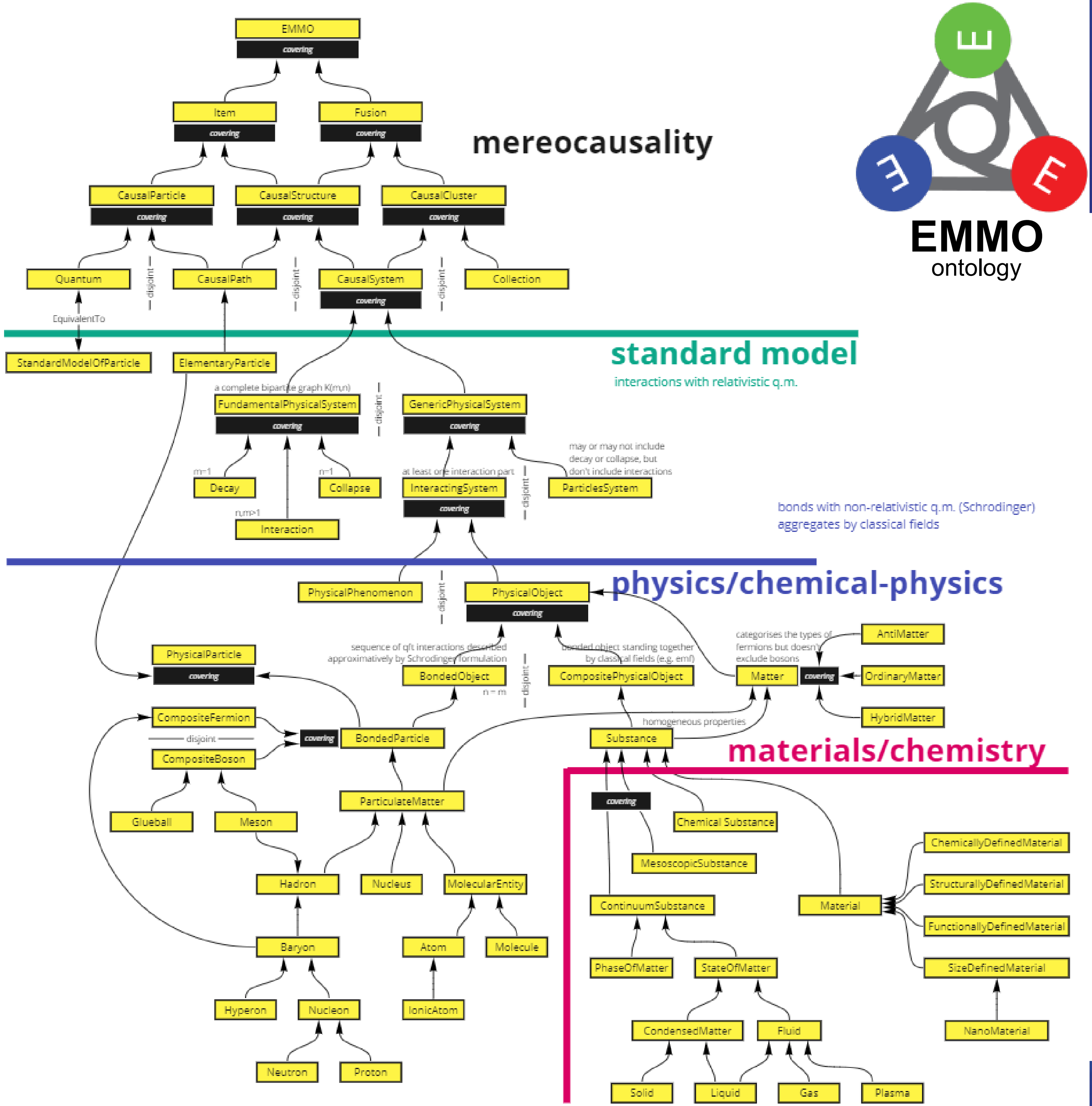
https://github.com/emmo-repo/EMMO/



EMMO structure

EMMO achieves complete **multiscale** representational capabilities with solid foundations in physics





FEMS 2025 EUROMAT

Ontologies Ecosystem in EMMO: Batteries

EMMO Top
(Mereocausality)

EMMO Perspectives
Persistence - Holistic - Contrast - Semiotics - ...

EMMO Reference
Standard Model/Physicalistic - Agency - Data/Information - Workflow -

EMMO Disciplines
Matter/Chemistry/Materials - Manufacturing - Metrology - Properties - Models - Maths - ....

Chemical Substance (battery) | Battery (whole) | Electrochemistry | CHAMEO | Battery Value Chain

CHEMS — Chemical Substance Domain Ontology
BATTERY — Battery Domain Ontology
ECHO — Electrochemistry Domain Ontology
CHAMEO

Example

Causal Structure

Process (Persistence)

Agent Driven Process, IntentionalProcess

Manufacturing, Materials Processing ISO and DIN standards

ElectroCoating

Plus alignment with: **Semantic Sensor Network Ontology**

OGC® Making location count. W3C®

References

[1] Battery Cell Assembly Twin (BatCAT) http://www.batcat.info/

[2] Elementary Multiperspective Material Ontology, https://github.com/emmo-repo/emmo

[3] Characterisation Methodology Domain Ontology (CHAMEO), https://github.com/emmo-repo/domain-characterisation-methodology

[4] Electrochemistry Ontology https://github.com/emmo-repo/domain-electrochemistry

[5] Battery Domain Ontology, https://emmo-repo.github.io/domain-battery/index.html

[6] Semantic Sensor Network Ontology, https://www.w3.org/TR/vocab-ssn/

[7] Chemical Substance Ontology, https://github.com/emmo-repo/domain-chemical-substance

[8] Battery Value Chain Ontology, https://github.com/Battery-Value-Chain-Ontology/ontology

BatCAT Battery Cell Assembly Twin

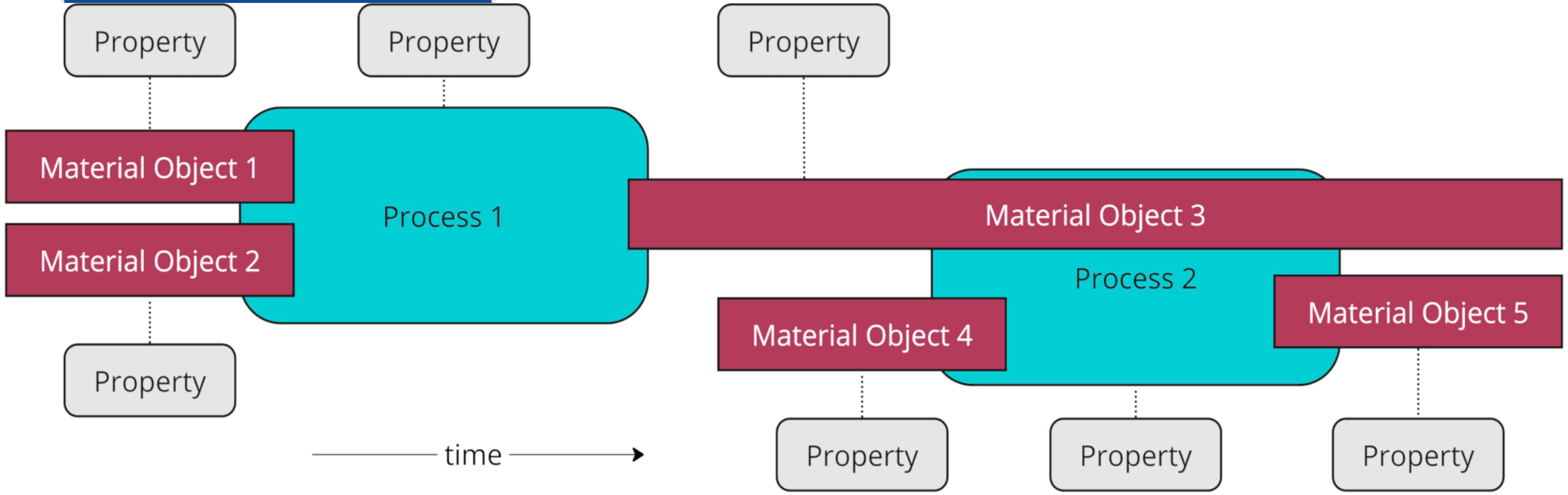
EMMO Developments to support wide use

- A stable, logically sound Upper and Middle level ontology
  - Not meant for general user application
  - Key concepts are ‘common sense’ to scientists (e.g. causality)
- Discipline ontologies (and taxonomies) as foundations for domains and applications, based on standards as much as possible.
- ELITE: A lightweight, user- and human-friendly version for applications development

EMMO-LITE Public

https://github.com/emmo-repo/EMMO-LITE

Representing manufacturing process in EMMO



Material, **Process**, and **Property** entities are described by EMMO classes. Industrial workflows are described as 4D connected processes and objects.

Post-launch analysis of pouch detergents

Fast analysis of large datasets to assess in-market dynamics across multiple countries.

Detergent Pouch Systems

Fast product development of laundry detergent pouch systems and tailoring to different customer segments.

Composite Prepregs

Achieve improved understanding and control the manufacturing process of bio-based composite prepregs.

Steel Section Mill

Deploy a coherent, seamless, and unified system covering all aspects of steel manufacturing.

Template to construct ontology of the application case

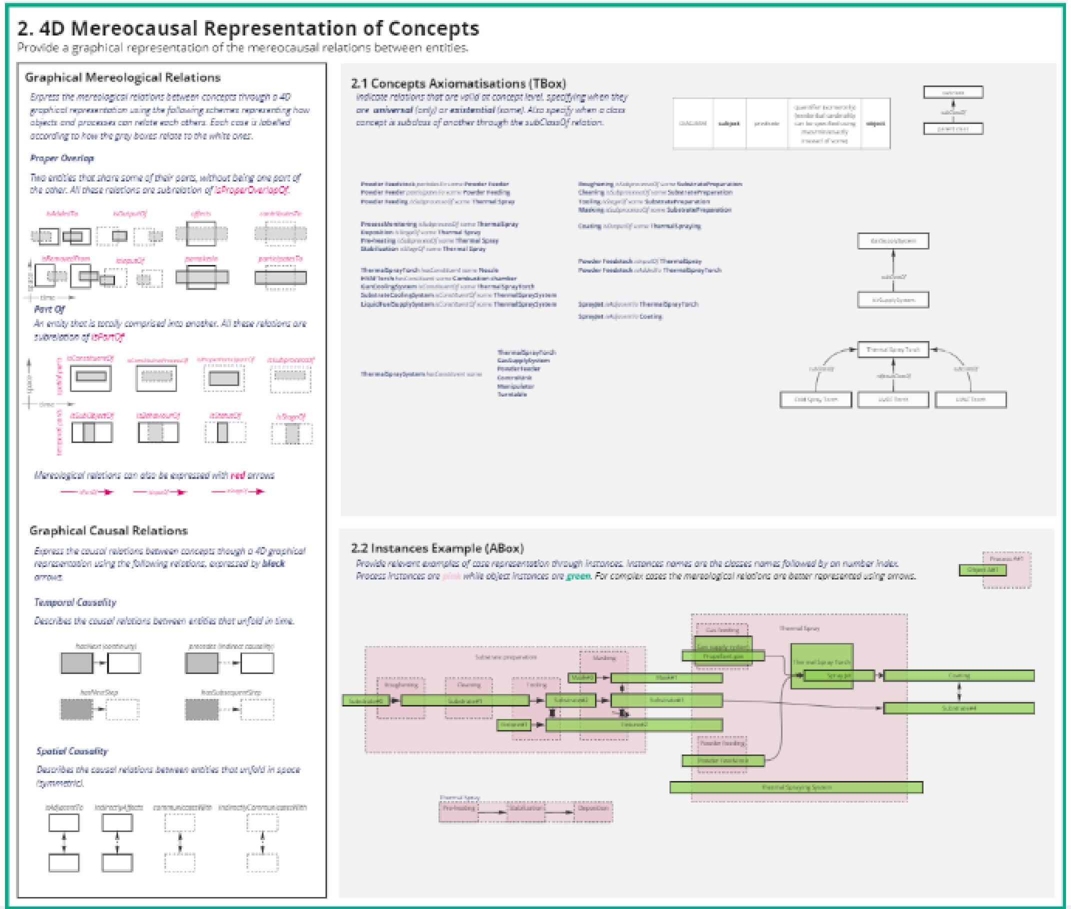
FEMS 2025 EUROMAT

ONTOTRANS

CoBRAIN

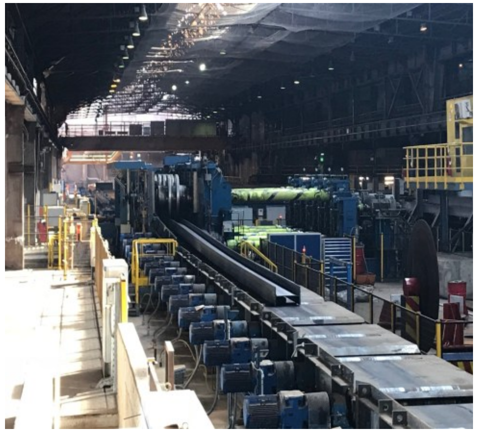
DESCRIBE

RELATE



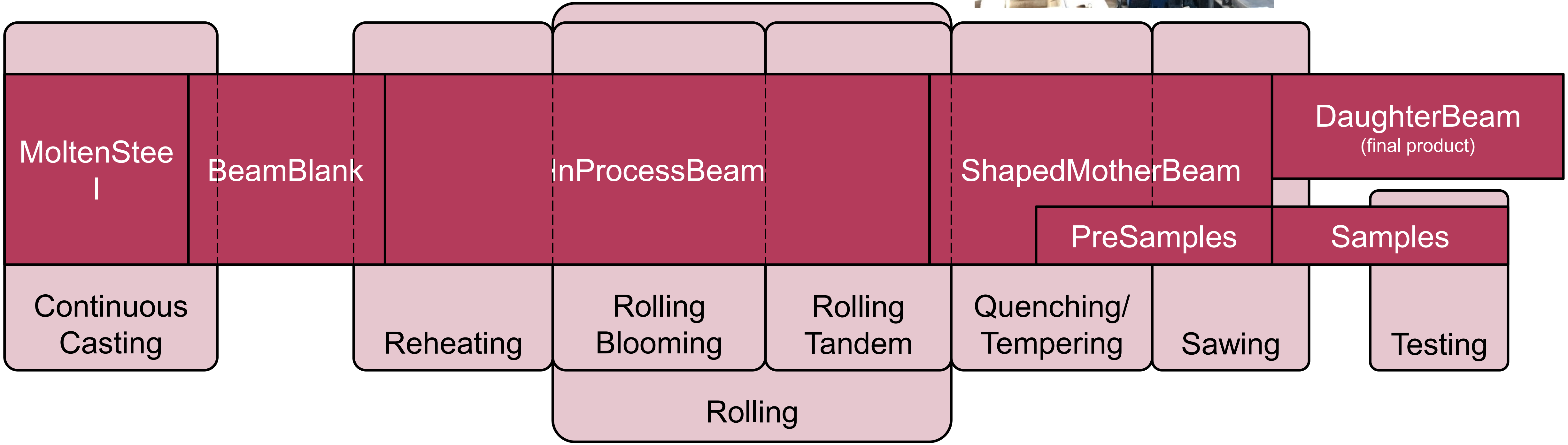
FEMS 2025 EUROMAT

Section Mill (Steel Rolling)



ONTOTRANS

ArcelorMittal



FEMS 2025 EUROMAT

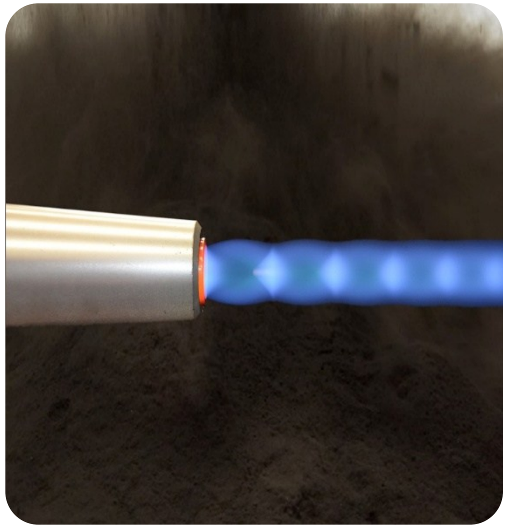
The CoBRAIN experience

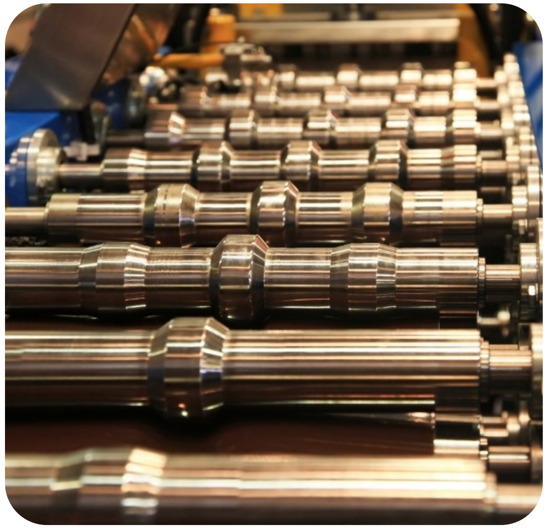
CoBRAIN

Funded by the European Union

Horizon Europe, GA no 101092211

INTEGRATED COMPUTATIONAL/EXPERIMENTAL MATERIAL ENGINEERING OF THERMAL SPRAY COATINGS





1. Planning: Data Management Plan

Focus on how to extract relevant information

Includes Tables for all datasets with fields (metadata) for each type of data!

Divided into **SHARED DATA** and **LOCAL DATA**:

SHARED DATA and METADATA Information that is uploaded in the CoBRAIN Knowledge Base: it has to be in the form of numerical or string values or its location specified using URIs or similar.

LOCAL DATA: a curator must be identified

Nanoindentation - High speed 3D mapping

SHARED DATA and METADATA

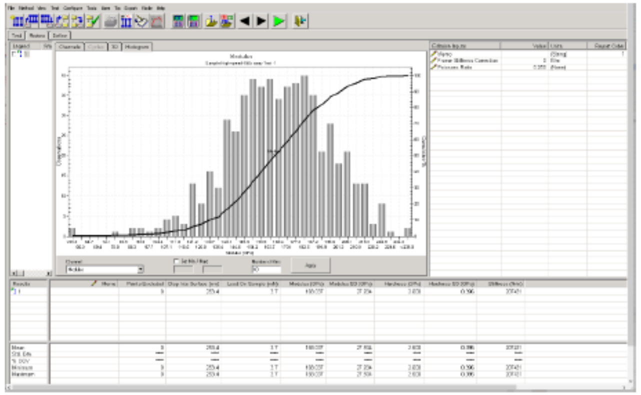
Field	Description
Sample ID	String: Sample name and identification number of test.
Sample description	String: sample composition and production information.
Test definition	String: Type of test and instrument name (i.e. Nanoindentation-High speed 3D mapping, MTS G200).
SOP file name	String: indication of the file name containing parameters of measurement (i.e. acquisition rate, max load, loading rate, approach speed etc.) .
Measurement Date Time	dd/mm/aaaa hh:mm:ss
Indenters	String: type of the indenter (i.e. Berkovich); serial number; material of the indenter
H, E maps	Data for 3D hardness and modulus maps (i.e. hardness value for x-y positions).

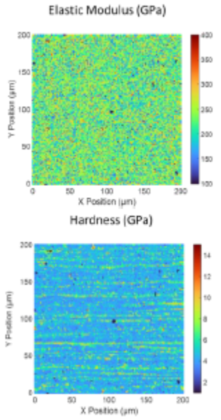
RDFS

LOCAL DATA

Field	Description
Sample ID	String: Sample name and identification number of test.
Measurement file	**.mss**, located in UNIROMA3's servers, contains all the information related to the measurement, and the complete raw results.
Indentation map	**.PNG**, located in UNIROMA3's servers, contains pictures of the indentation maps, (i.e. applied load [mN] vs penetration depth [nm]).
Indentation Data	**.dat**, located in UNIROMA3's servers, contains all data from raw to final to obtain 3D map (i.e. applied load [mN] vs penetration depth [nm]).

non RDFS

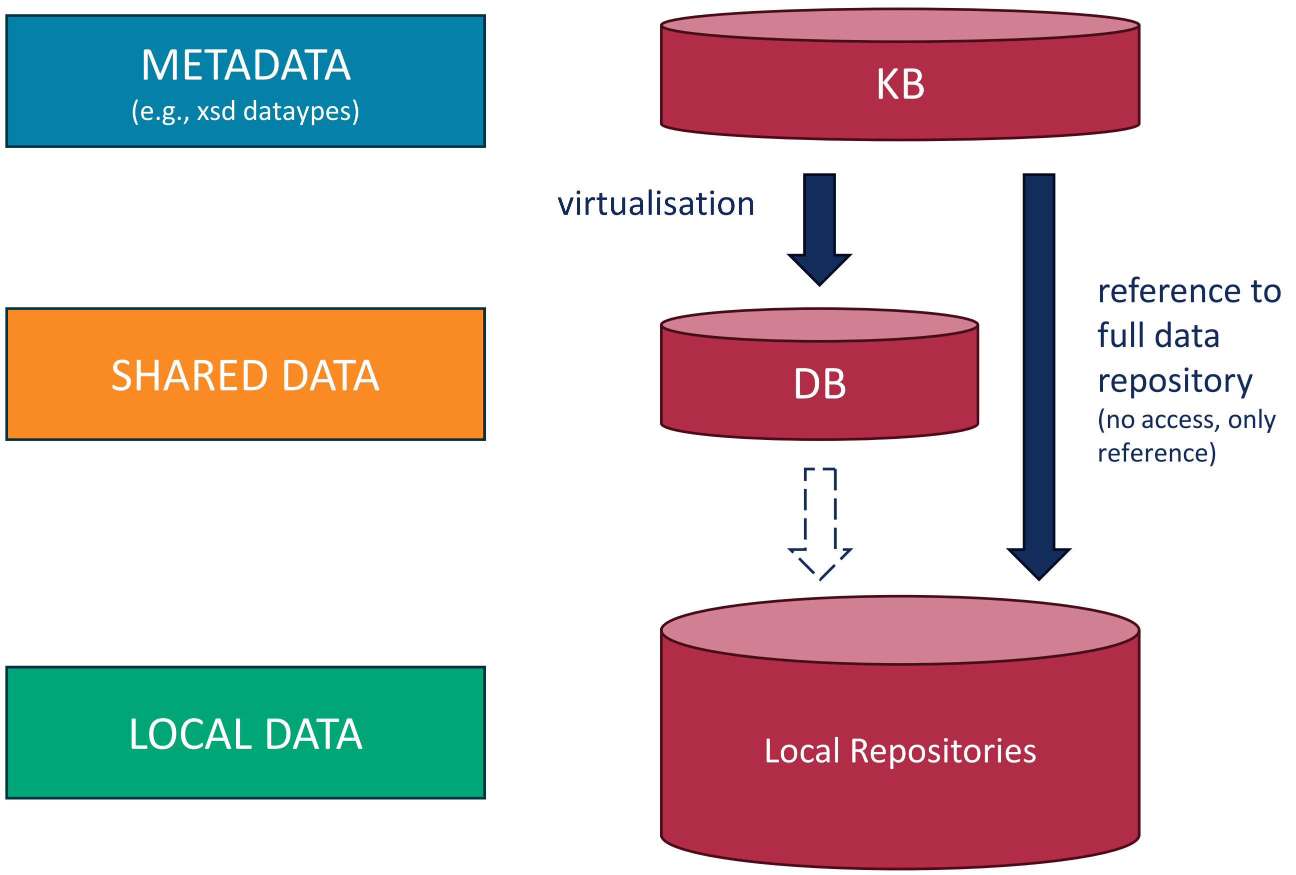




1. Planning: knowledge base architecture

Co-BRAIN

Original Architecture



SHARED DATA and METADATA

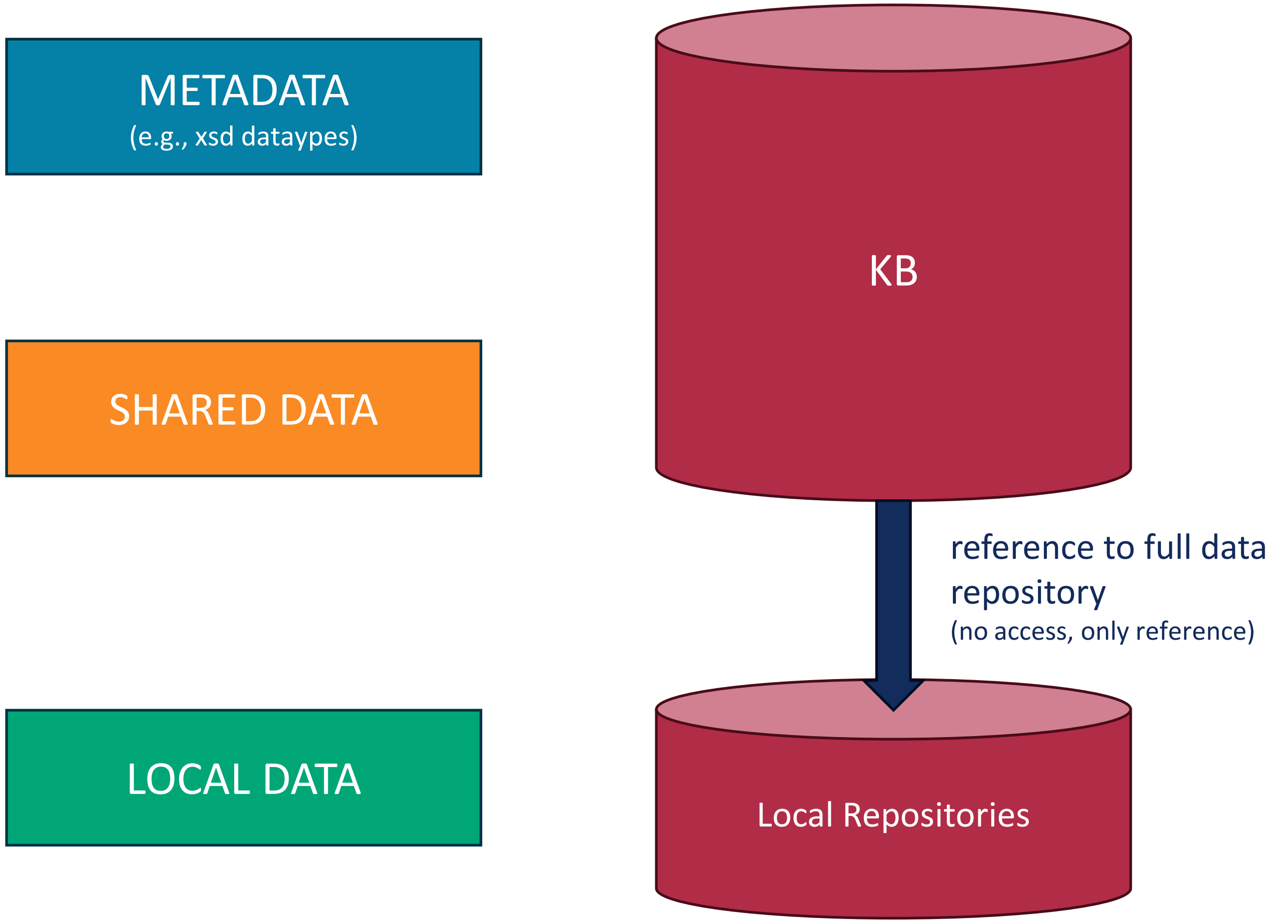
Field	Description
Sample ID	String: Sample name and identification number of test.
Sample description	String: sample composition and production information.
Test definition	String: Type of test and instrument name (i.e. Nanoindentation-Pillar splitting, MTS G200).
SOP file name	String: indication of the file name containing parameters of measurement (i.e. acquisition rate, max load, loading rate, approach speed etc.) .
Measurement Date Time	dd/mm/aaaa hh:mm:ss
Indenters	String: type of the indenter (i.e. Berkovich); serial number; material of the indenter.
Curves	Indentation curve: Load applied [N] vs penetration depth [m].
P_C	**mN,** the value of the critical load, identified from Indentation curve: (Load applied [N] vs penetration depth).
K_{IC}	**MPa√m,** the value of the fracture toughness calculated from the value of the critical load.

LOCAL DATA

Field	Description
Sample ID	String: Sample name and identification number of test.
Measurement file	**.mss**, located in UNIROMA3's servers, contains all the information related to the measurement and the complete results.
Indentation Data	**.xls**, located in UNIROMA3's servers, contains data, exported from .mss, needed for fracture toughness calculation (i.e. applied load [mN] vs penetration depth [nm]).

2. Real life: No local db management

Unfortunately, local data management in the form of local query-able databases (e.g., MongoDB, SQL) as foreseen in the GA was not a realistic possibility for partners… so we dropped virtualisation



SHARED DATA and METADATA

Field	Description
Sample ID	String: Sample name and identification number of test.
Sample description	String: sample composition and production information.
Test definition	String: Type of test and instrument name (i.e. Nanoindentation-Pillar splitting, MTS G200).
SOP file name	String: indication of the file name containing parameters of measurement (i.e. acquisition rate, max load, loading rate, approach speed etc.) .
Measurement Date Time	dd/mm/aaaa hh:mm:ss
Indenters	String: type of the indenter (i.e. Berkovich); serial number; material of the indenter.
Curves	Indentation curve: Load applied [N] vs penetration depth [m].
P_C	**mN,** the value of the critical load, identified from Indentation curve: (Load applied [N] vs penetration depth).
K_{IC}	**MPa√m,** the value of the fracture toughness calculated from the value of the critical load.

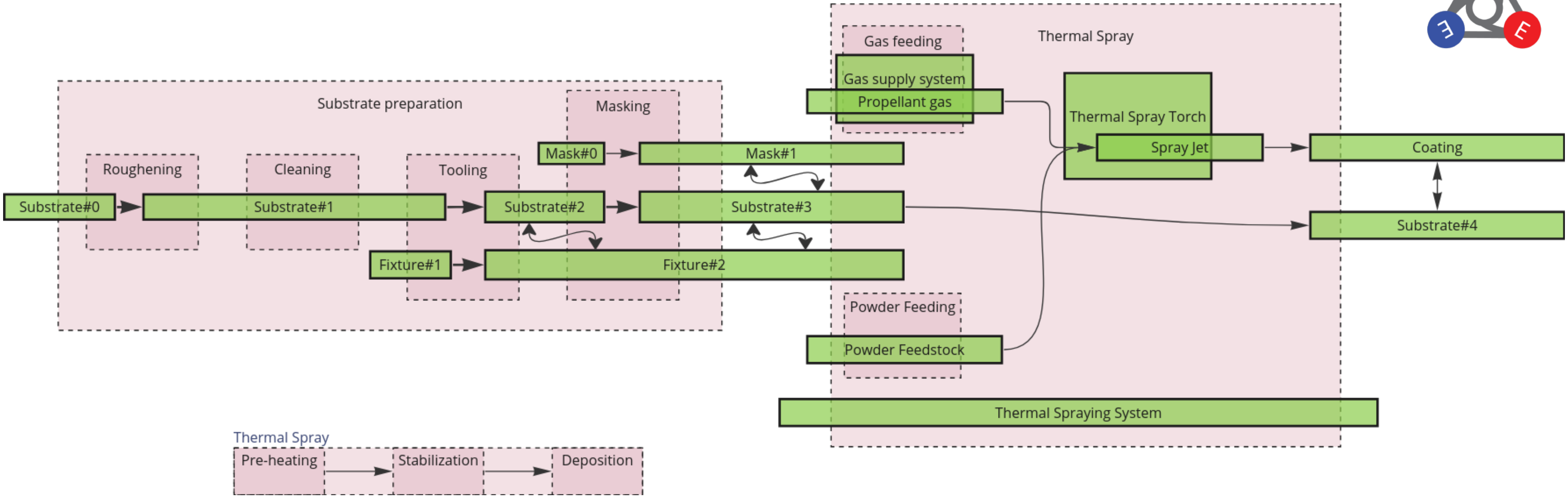
LOCAL DATA

Field	Description
Sample ID	String: Sample name and identification number of test.
Measurement file	**.mss**, located in UNIROMA3's servers, contains all the information related to the measurement and the complete results.
Indentation Data	**.xls**, located in UNIROMA3's servers, contains data, exported from .mss, needed for fracture toughness calculation (i.e. applied load [mN] vs penetration depth [nm]).

1. Planning: data/entities model of the applications

Co-BRAIN

4D conceptual representation of a Thermal Spraying process in CoBRAIN project connecting all the relevant entities using EMMO mereocausality relations.



1. Planning: Formalisation in an ontology

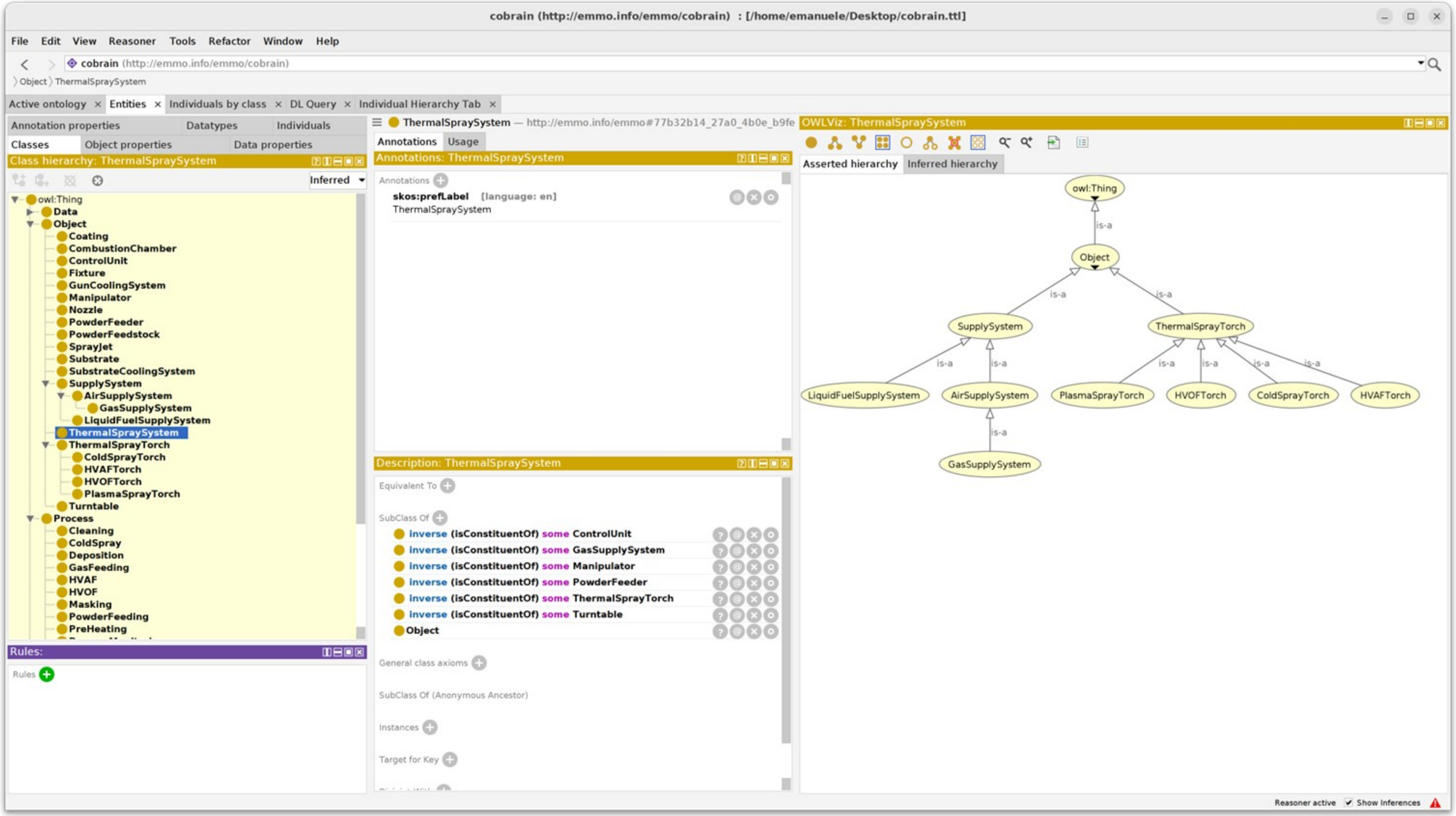
Co-BRAIN

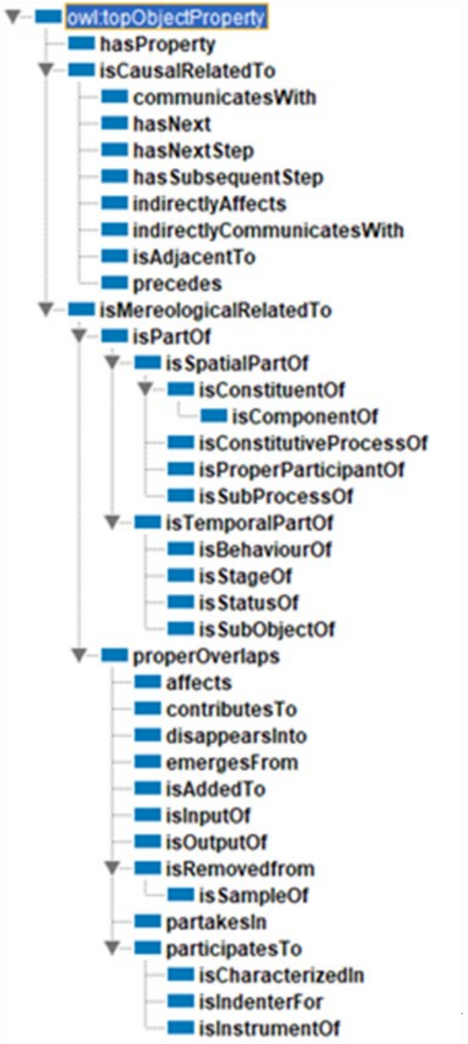
CoBRAIN Thermal Spraying Ontology

https://github.com/cobrain-project

EMMO-LITE Public

https://github.com/emmo-repo/EMMO-LITE





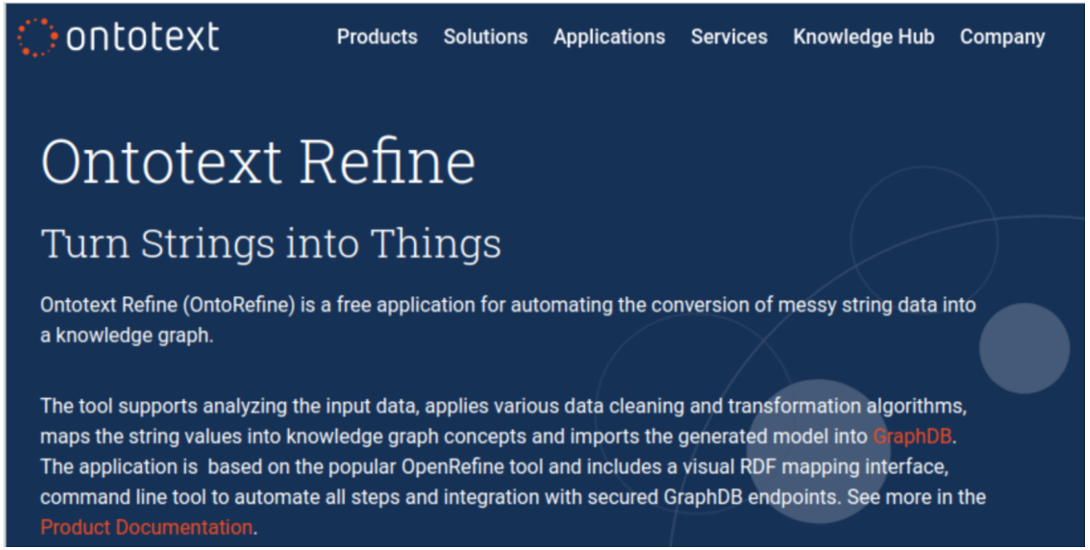
1. Planning: Software

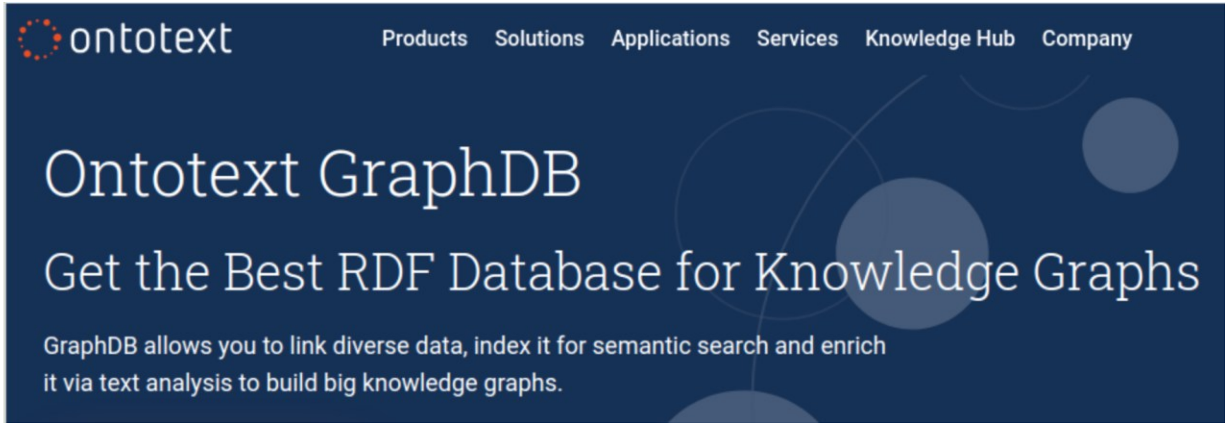
Don't reinvent the wheel

Use existing commercial tools with free-to-use licensing option

Adhere to W3C standards

Benefit from high level of maturity.





2. Acquiring: Data in Excel tables

Issue: **partner specific methods and formats**.

- Heterogenous symbols: “º” (U+00BA) vs “°” (U+00B0)
- Arbitrary names, e.g. AlCrFeMnNi referred as Al14.
- Sometimes the lotNr is called Batch Nr…
- Format inconsistencies

Solution:

- Provide **constrained templates** for data collection
- **Refactoring** all collected data centrally (by Univ Bologna)

METADATA												DATA		
Sample				Test					Indenter					
Sample ID	Composition	Production Information	Origin ID	Test ID	Test Type	Instrument ID	SOP File	DateTime	Type	S/N	Material	Indentation Curve	Pc	KIC
FC0Oq8	[[Cu,Zn],[65,35]]	Some text about sample production	yQvzd7	h2baA2	PillarSplitting	MTSG200-1	file://emanuele@kant.unibo.it:22/home/emanuele/file3.txt	12/3/2023 11:54	Berkovich	XlUQ32	W	[[1,2,3],[1,2,3]]	12	34
TatORe	[[Cu,65],[Zn,35]]	Some other text about sample production	yQvzd7	ap6eUT	PillarSplitting	MTSG200-1	file://emanuele@kant.unibo.it:22/home/emanuele/file5.txt	14/3/2023 12:54:00 AM	Berkovich	XlUQ32	W	[[1,3,4],[2,1,5]]	5	3

Future projects: implement **a full framework for assisted/automated data collection** for all datasets, providing a **built-in semantic enhancement of data**

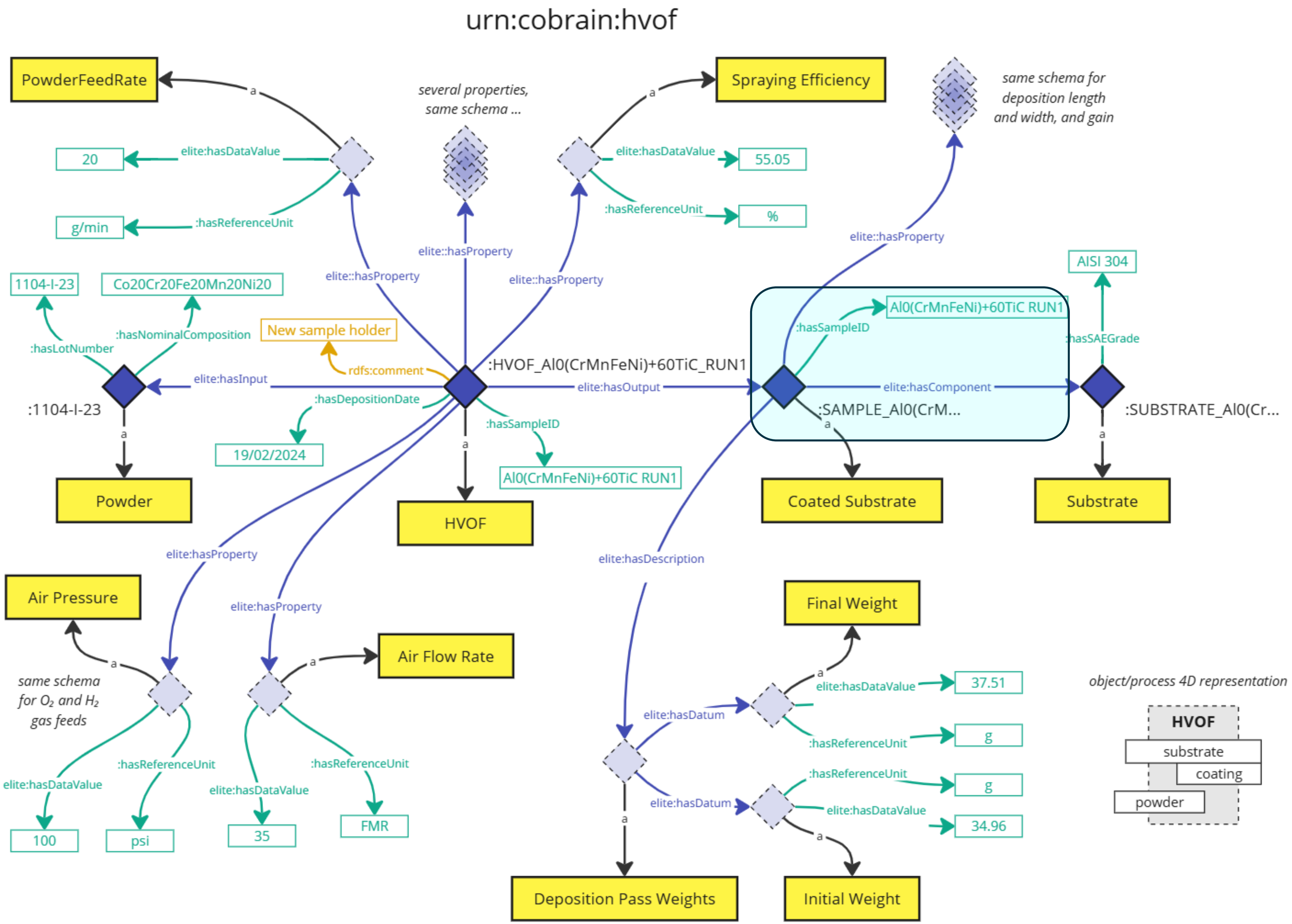
Requires substantial effort in first phase of project, typically not done….

3. Processing: Map data to the interconnected entities

Graphical Approach to Conceptual Mapping

Partners collaborated with UNIBO to build the conceptual mapping between the Excel file columns and the ontological concepts, including relations.

UMR HVOF Thermal Spraying Logbook

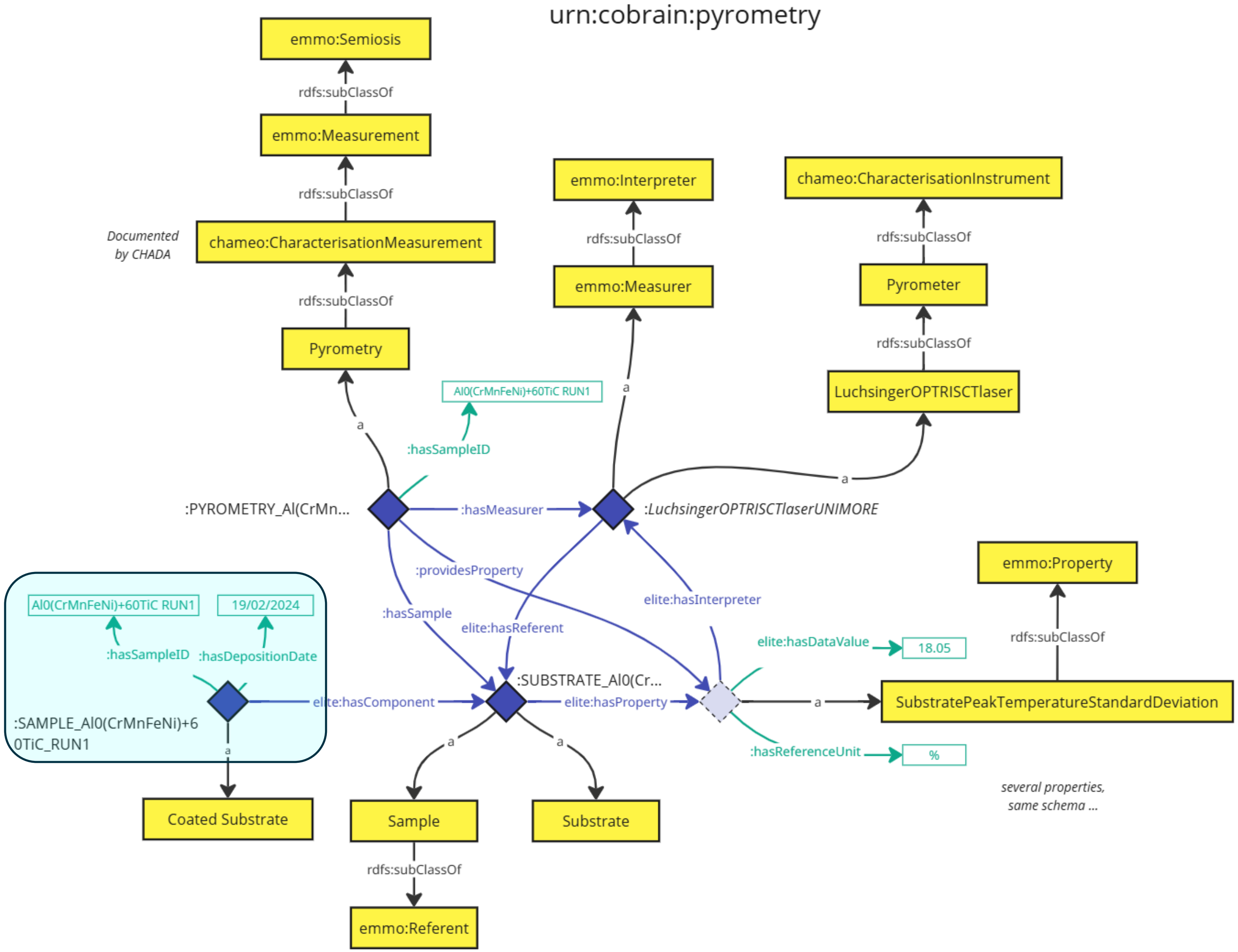


3. Processing: Map data to the interconnected entities

Graphical Approach to Conceptual Mapping

Partners collaborated with UNIBO to build the conceptual mapping between the Excel file columns and the ontological concepts, including relations.

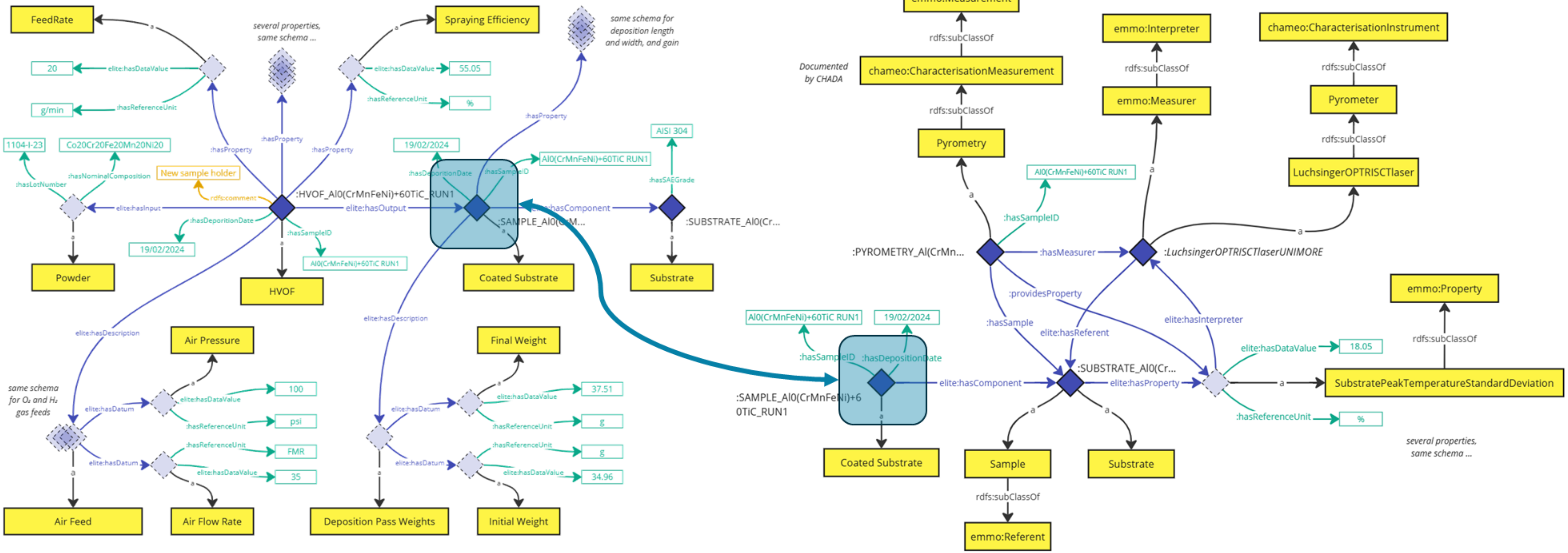
Optical pyrometer Luchsinger OPTRIS CTlaser



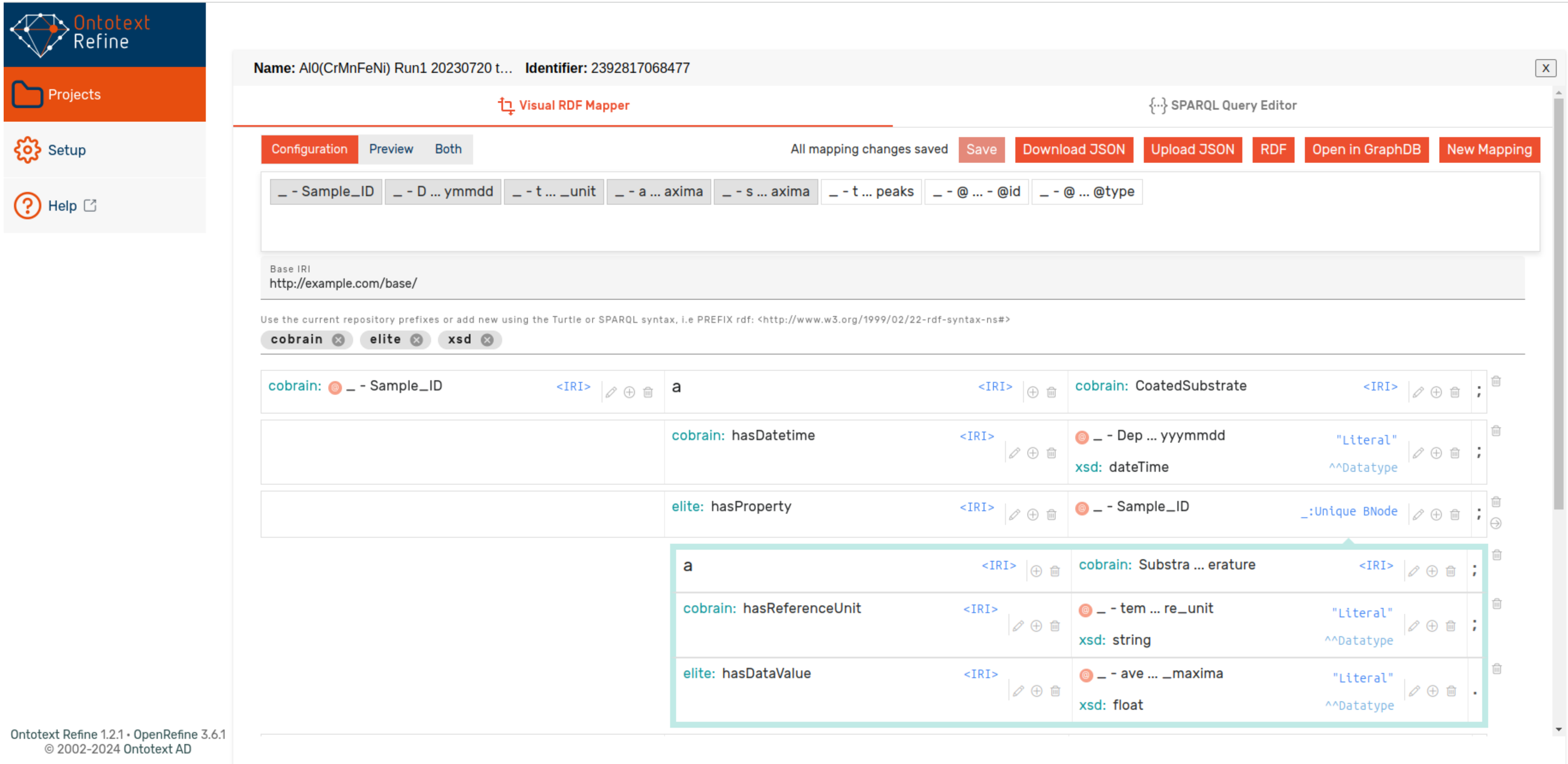
3. Processing: interoperability

UMR HVOF Thermal Spraying Logbook

Optical pyrometer Luchsinger OPTRIS CTlaser



3. Processing: Mapping with OntoRefine



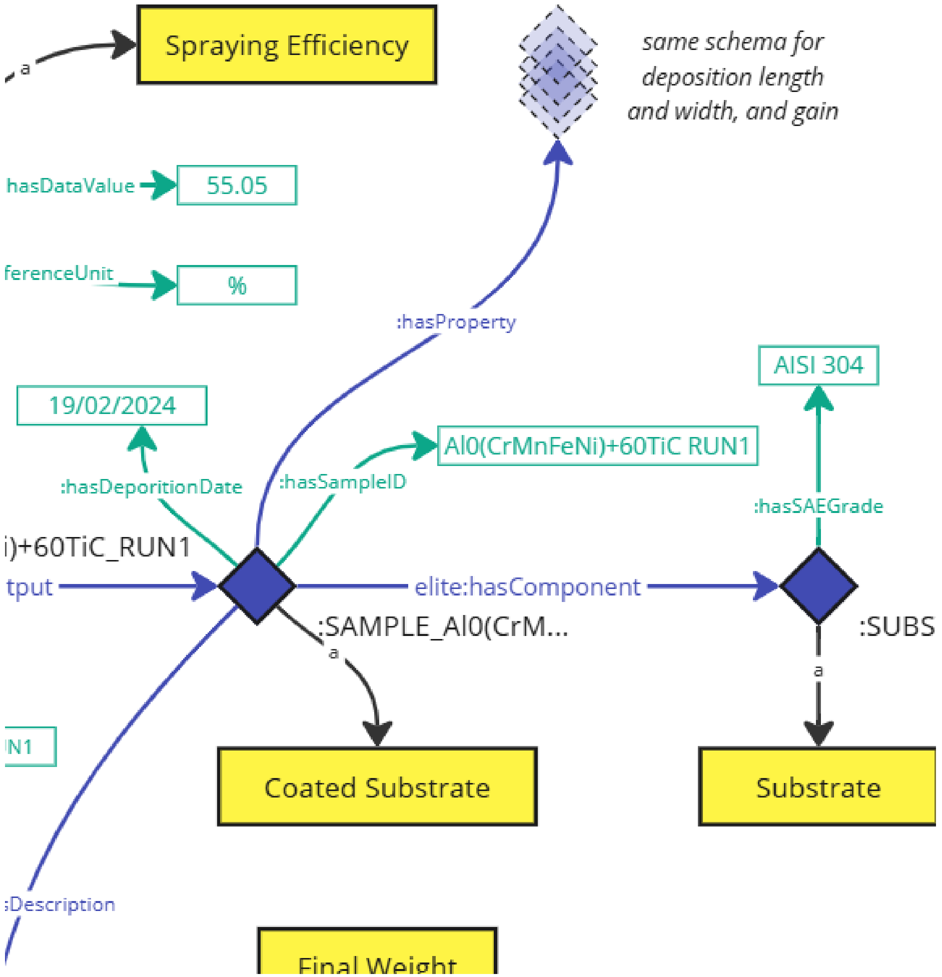
3. Processing: practical issues

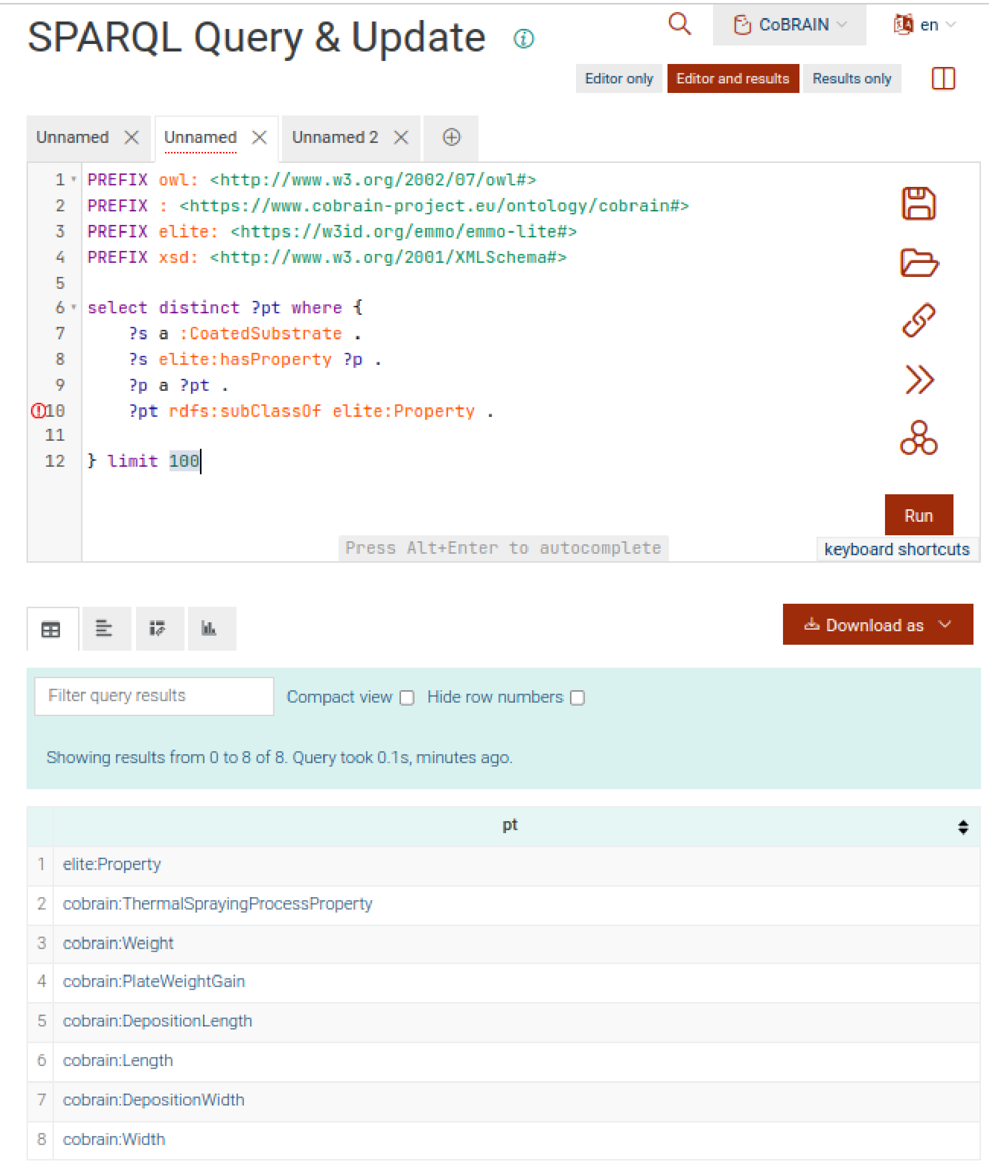
"Dirty data"

- **Unprecise IDs** (e.g., case change, spaces, use of reserved character)
- **Non-Unique IDs**: same ID used for e.g., deposition, sample, substrate…
- **Multiple intended interpretation for the same cell-value**, leading to an excessive blank-node rich representation
- **Too many degree of freedom** for the user to fill the sheets, leading to inconsistencies and scattered text

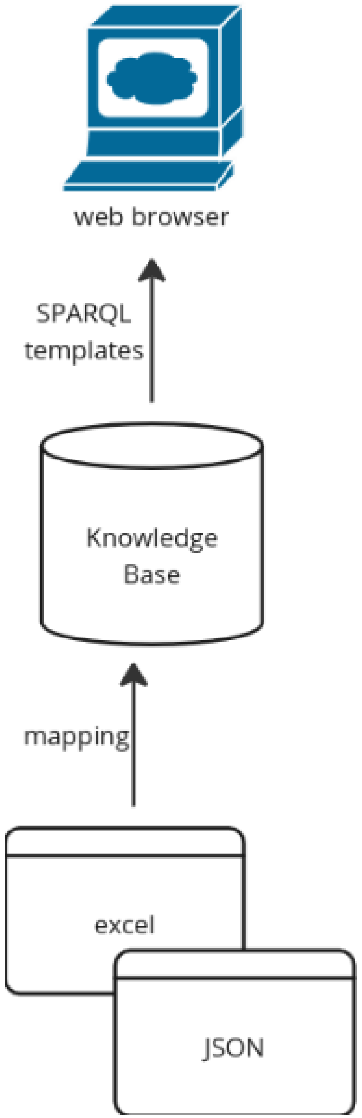
4. Analysing: support users with SPARQL queries

Conceptual schemes are paramount to elucidate the **CoBRAIN ontology data model** and **concept relations**, in order to **enable non-expert users** to generate the **correct SPARQL query** for their needs.

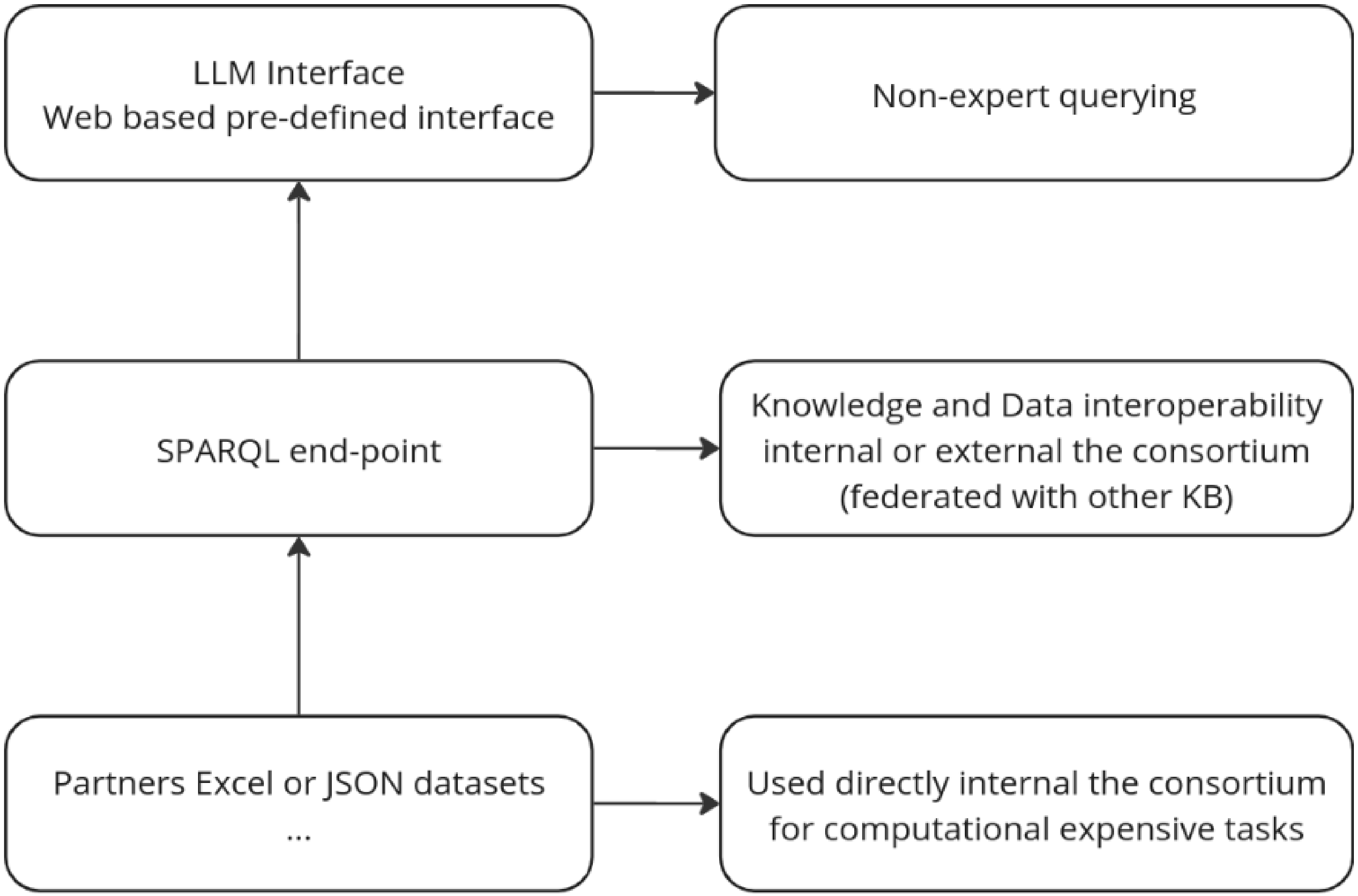




5. Sharing



web browser
SPARQL templates
Knowledge Base
mapping
excel
JSON



LLM Interface
Web based pre-defined interface
Non-expert querying
SPARQL end-point
Knowledge and Data interoperability internal or external the consortium (federated with other KB)
Partners Excel or JSON datasets
...
Used directly internal the consortium for computational expensive tasks

6. Outcomes

- A **federated interoperable database** for the CoBRAIN partners that can be used also after the project ends
- **FAIR solution** for the dissemination and exploitation of project data
- Support the creation of a **SDSS** (Sustainable Decision Support System)

The **CoBRAIN ontology for Thermal Spraying** can be used by other research groups to make their data interoperable with CoBRAIN.

CoBRAIN Status

- A **fully operable methodology** for the KB, based on the free version of a well-known commercial triplestore (GraphDB) and tools (OntoRefine)
- **Easy to deploy and share** in other frameworks (the overall KB can be shared as a single file, ontology included)
- The TBox is based on **EMMO-LITE** and CoBRAIN ontology, usable with OWL 2 RL/QL/EL profiles
- **Established a low-mid level framework for FAIR data in the field of thermal spraying of materials**
- Proposed **user-friendly methodologies** for data collection in **experimental environment**
- Provide a **scalable and federated** environment for knowledge management that can be expanded in the future with external DB (e.g. Mongo DB)
- Set of **semantically connected datasets** for interoperable SPARQL queries
- **Knowledge Base** deployed in a test server.

Lessons Learned

CoBRAIN

- Cannot assume any data generated in the project are coherent and well-structured.

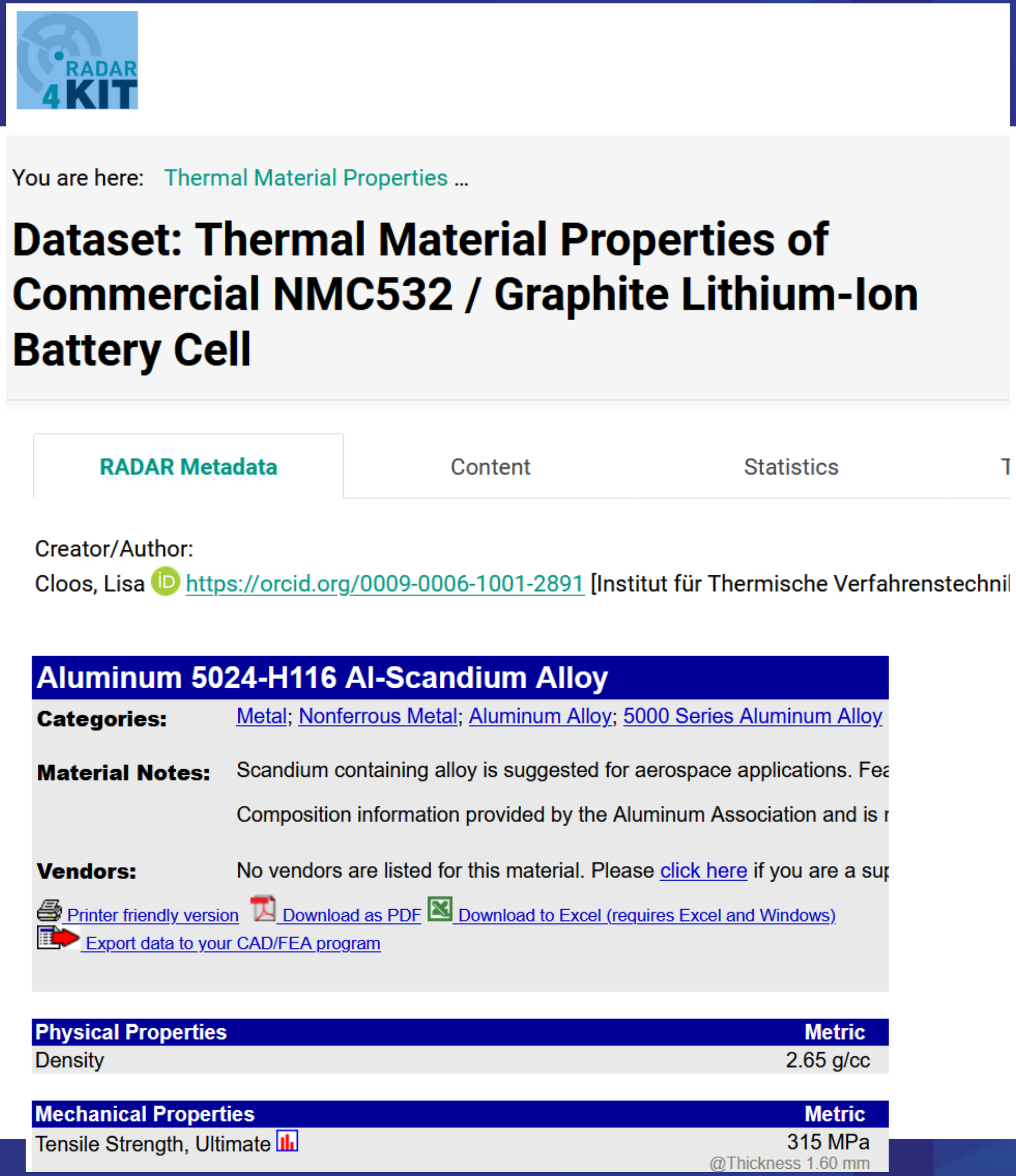
- **Data Management /stewardship at partner/local level requires significantly more resources**. Improvement of data management practices is beneficial in any case.

- Accommodating different data storage practices is complex.

Requirements

- For Datasets:
  - cataloguing based on generic and materials specific metadata; Materials DCAT-AP
  - Controlled Vocabularies

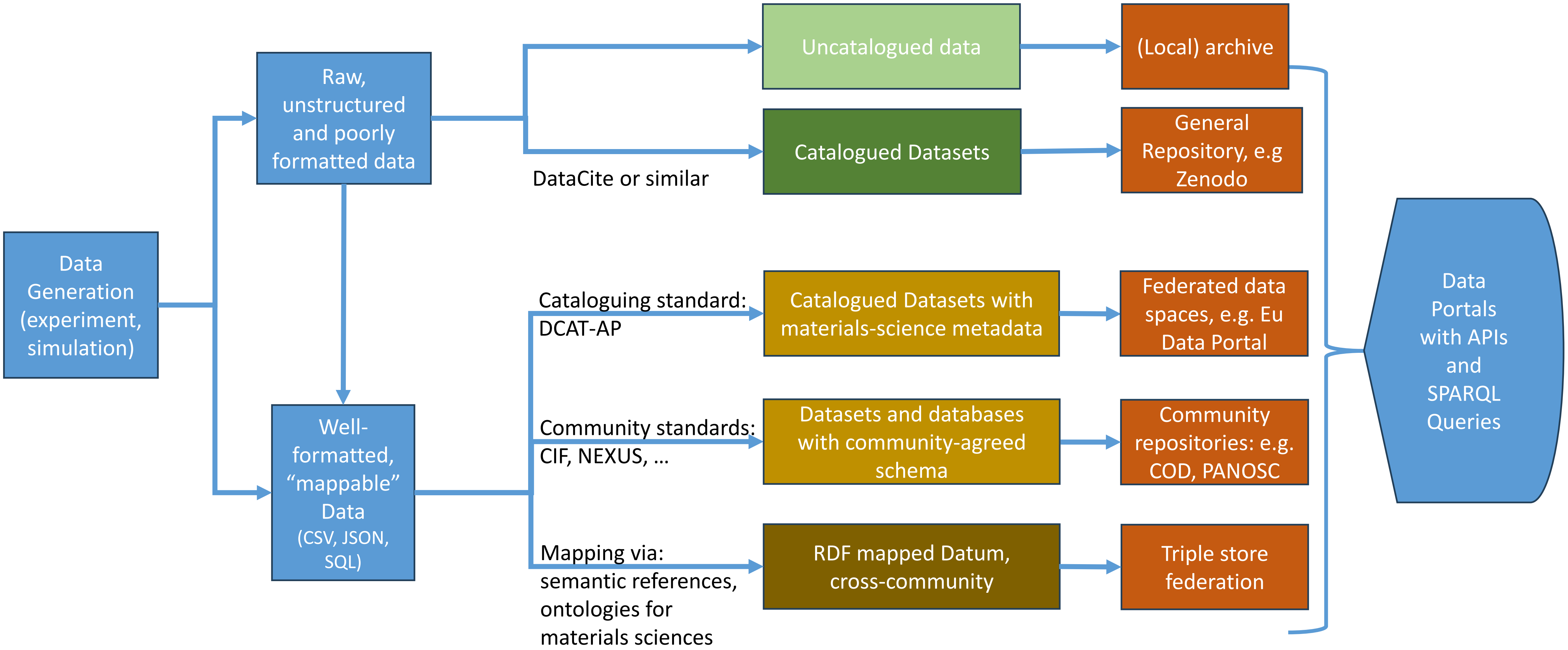
- For full data detail integration
  - RDF mapping with a coherent set of ontologies
  - Support data federation, federated queries



https://www.matweb.com/

Data integration architecture

Inspired by CONNECT-NM
https://www.connect-nm.eu/



Outlook: Materials Commons for Europe

https://materialscommons4.eu/

MaterialsCommons4EU - Driving Digital Innovation in Materials Science within Europe
Cross-sector collaboration via **a federated digital infrastructure.**
Building on strong synergies with key initiatives such as Material Digital, DIADEM, MaX, CaPeX, and numerous other national platforms.

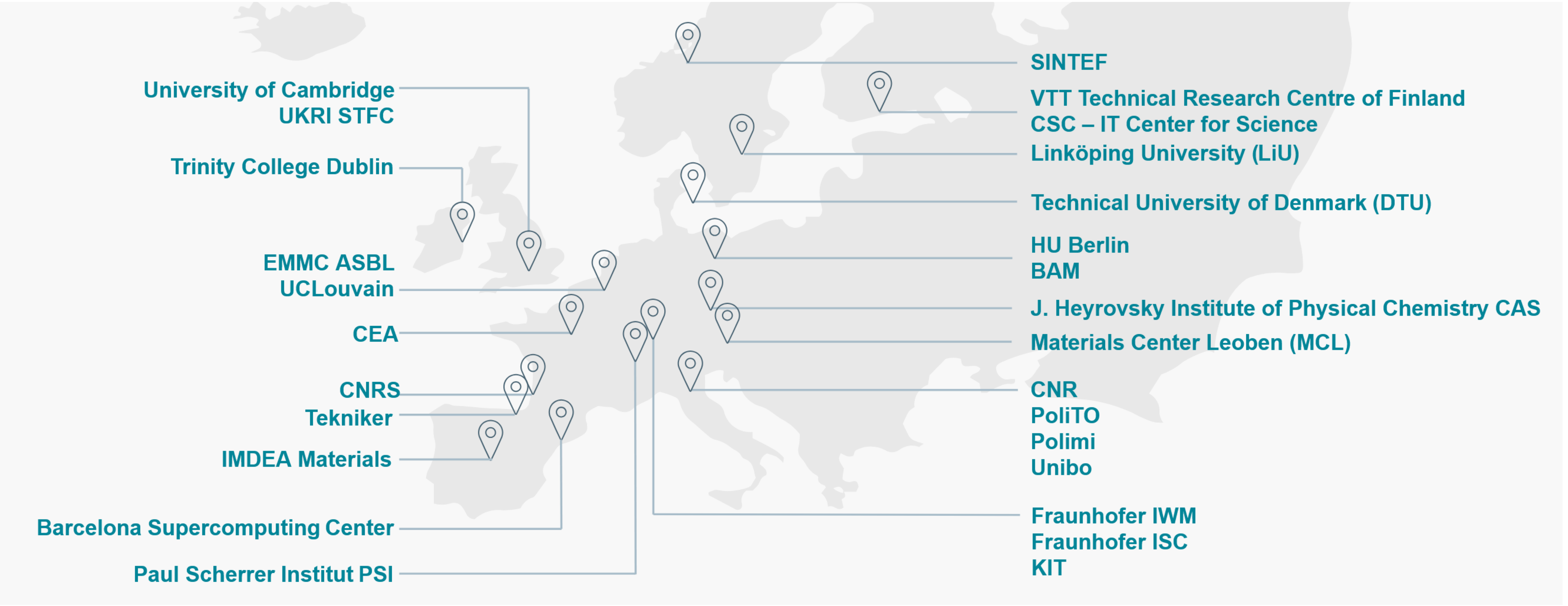
General information

Programme
Horizon Europe (HORIZON)

Call
INDUSTRY (HORIZON-CL4-2025-01)

Type of action
HORIZON-IA HORIZON Innovation Actions

Deadline model
single-stage



FEMS 2025 EUROMAT

Acknowledgements

- Emmanuele Ghedini, University of Bologna, and the CoBrain project
  - funding from the European Commission under the European Union’s Horizon Research and Innovation programme, GA no 101092211
- Ilchat Sabirov, IMDEA and AID4GREENEST colleagues
  - funding from the European Commission under the European Union’s Horizon Research and Innovation programme, GA no 101091912
- Goldbeck Consulting Ltd and the BatCAT project
  - funding from the European Commission under the European Union’s Horizon Research and Innovation programme and by the UK Research and Innovation (UKRI) under the UK government’s Horizon Europe funding guarantee, GA no 10091190

EMMC

CoBRAIN

aid4 GREENEST

GOLDBECK CONSULTING

BatCAT Battery Cell Assembly Twin

Thank you!

Abstract

EMMC is a community of stakeholders covering all types of materials models, computational methods and applications. Improved methods integration and digitalisation are of particular relevance to advancing and accelerating materials design in a way that takes multiple dimensions of properties, performance and circularity into account. To that end, EMMC has been facilitating workflow integration, data documentation and semantic interoperability. Multiple projects have been collaborating under the EMMC banner in the development of the EMMO ontology for applied sciences and manufacturing. EMMO provides a strong semantic framework for all the datasets generated by experimental, modelling and characterisation activities. We give an overview of pertinent EMMO ontologies, including characterisation, manufacturing and their industrial applications. In more detail, we highlight the CoBRAIN project, demonstrating the impact of integrating computational and experimental data through semantic interoperability. It enables the development of intelligent tools that (a) propose novel materials from the class of high entropy alloy-based “hardmetals” for direct deposition by Thermal Spraying techniques, and (b) are capable to estimate their impact on the economy and the environment.

Development of new materials with high wear and good corrosion resistance is usually based on experimental trial-and-error and operational feedback. Combining data coming from modelling and experimental approaches provides an advancement but is limited by the fact that the final coating properties depend on too many materials and process factors. Hence obtaining reliable results from physical modelling as well as experimentation are in conflict with the response time and cost containment required for industrial innovation. Here, the application of modern AI techniques can speed up the innovation process when used as complementary tool to existing experimental, modelling and characterisation activities. It can also enhance decision systems integrating all aspects of the materials and production process, taking account of economic and environmental factors.

In order to facilitate such a system, all the CoBRAIN generated data are collected into a single end-point via a full conceptual model based on the EMMO ontology. Datasets from the various areas of expertise within the project have been integrated thanks to the interoperability provided by the EMMO framework and made available to the users through a triple store Knowledge Base. Besides dramatically improving the FAIR level of the project results, it facilitates the creation of meaningful and coherent dataset for AI training and verification.

Acknowledgement: The CoBRAIN project received funding by the EU Horizon Europe research and innovation programme under GA N. 101092211.