INTEGRATED COMPUTATIONAL/EXPERIMENTAL MATERIAL ENGINEERING OF THERMAL SPRAY COATINGS

EMMO for Manufacturing: the CoBRAIN Knowledge Base for Thermal Spraying Process, Modelling and Characterisation

MatCHMaker @ EMMC 2025 Satellite Workshop

7th April 2025, TU Wien, Vienna (AT)

Alma Mater Studiorum – Università di Bologna
Department of Industrial Engineering (DIN)
Bologna (Italy)

Presenters:

Emanuele Ghedini

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.

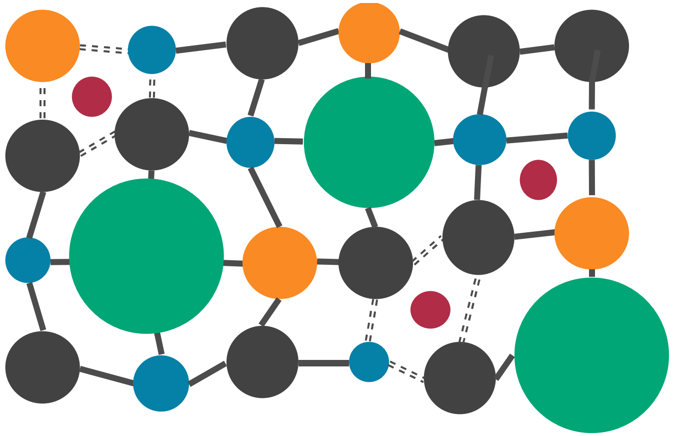
INTEGRATED COMPUTATIONAL/EXPERIMENTAL MATERIAL ENGINEERING OF THERMAL SPRAY COATINGS

- **Protective coatings** against wear and corrosion play a critical role in strategic industrial fields
- Extant technologies have considerable 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 (CRMs) – Co, W

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

The Project

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

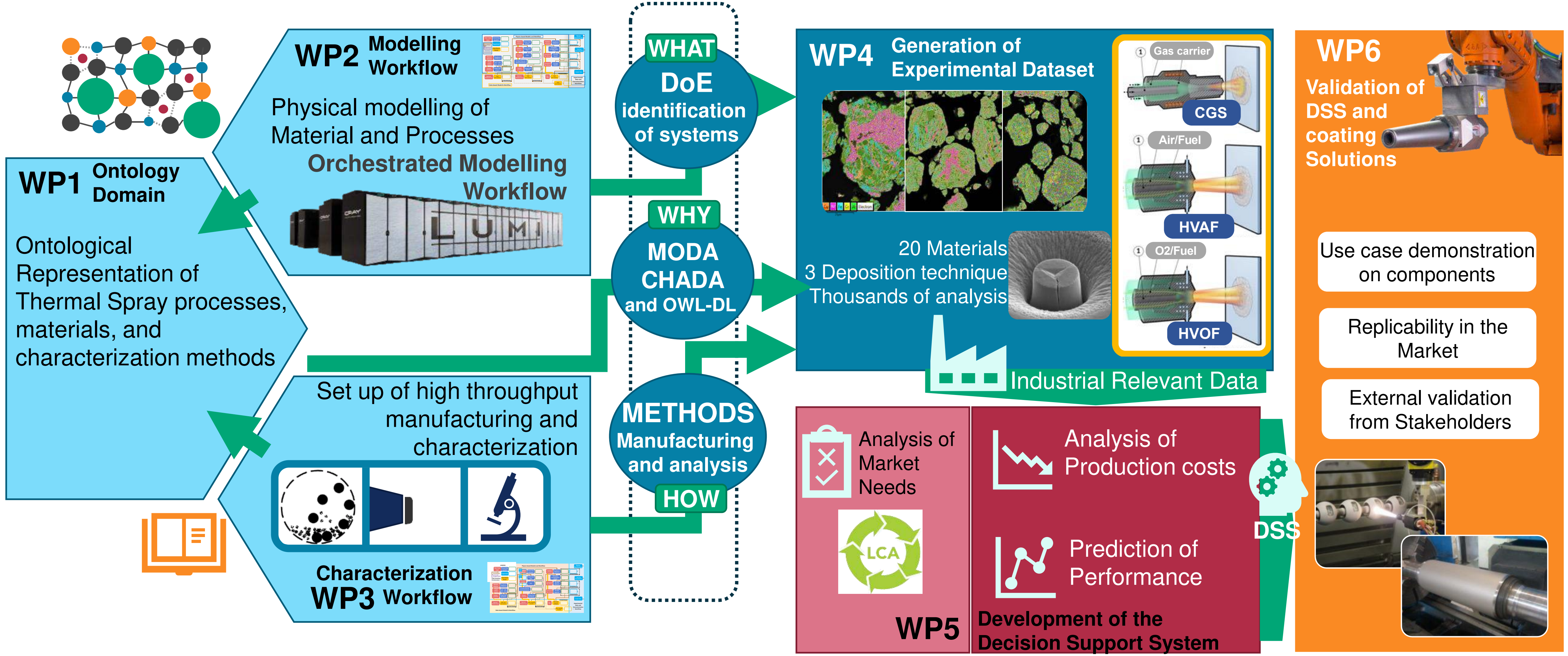
126 equiatomic combinations of 4 or 5 elements

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 stabilizing effect of the large entropy of a multi-element random solid solution ($\Delta G = \Delta H - T \cdot \Delta S$)
- **Lattice strain effect**
The mismatch in atomic radii result in lattice distortions that increase strength
- **Sluggish diffusion effect**
The distorted lattice hinders long-range diffusion, which also results in better high-temperature stability
- **“Cocktail” effect**
Non-linearities and unexpected synergistic effects may sometimes yield somewhat unpredictable outcomes

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

The Project


WP2 Modelling Workflow
Physical modelling of Material and Processes
Orchestrated Modelling Workflow
WP1 Ontology Domain
Ontological Representation of Thermal Spray processes, materials, and characterization methods
Set up of high throughput manufacturing and characterization
Characterization WP3 Workflow
WHAT
DoE identification of systems
WHY
MODA CHADA and OWL-DL
METHODS
Manufacturing and analysis
HOW
WP4 Generation of Experimental Dataset
20 Materials
3 Deposition technique
Thousands of analysis
Gas carrier
CGS
Air/Fuel
HVAF
O2/Fuel
HVOF
Industrial Relevant Data
Analysis of Market Needs
LCA
Analysis of Production costs
Prediction of Performance
WP5
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

1. Planning

o Did they plan a priory which data they needed, and how to extract knowledge from it?
o Did they work with Data Management Plans?

Co-BRAIN

CoBRAIN Data Management Plan

All datasets from **experiments**, **modelling** and **characterisation**, are documented in the DMP.

The tables are 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 (rdfs:literal). For data which are not in the form of rdfs shareable values (i.e., images, XRD plots) quantitative values must be provided and/or its location specified using URIs or similar.

LOCAL DATA Information that is kept at partner premises: for these data a curator must be identified, and it is specified whether the data are openly shareable on request or confidential.

Nanoindentation - High speed 3D mapping

SHARED DATA and METADATA

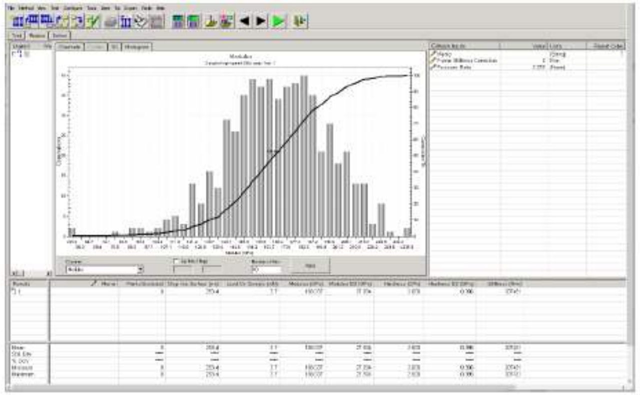
Field	Description
Sample ID	String: Sample name and identification number of test.
Sample description	String: sample composition and production information.
Test definition	String: Type of test and instrument name (i.e. Nanoindentation-High speed 3D mapping, MTS G200).
SOP file name	String: indication 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).

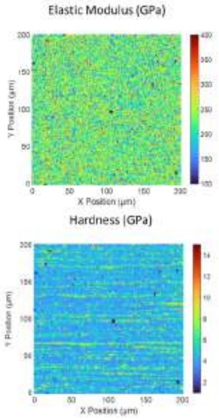
RDFS

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

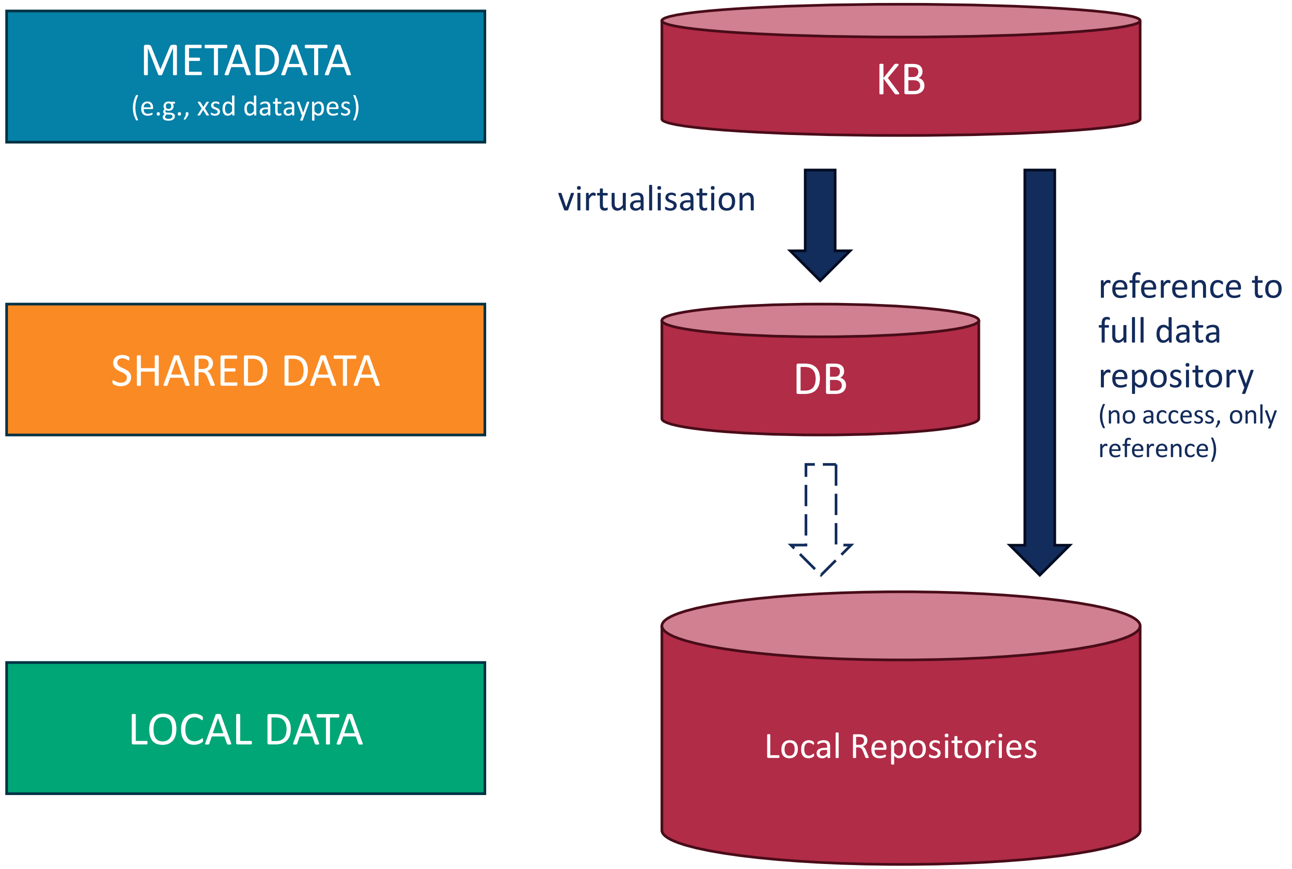




1. Planning

- Did they plan a priory which data they needed, and how to extract knowledge from it?
- Did they work with Data Management Plans?

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.

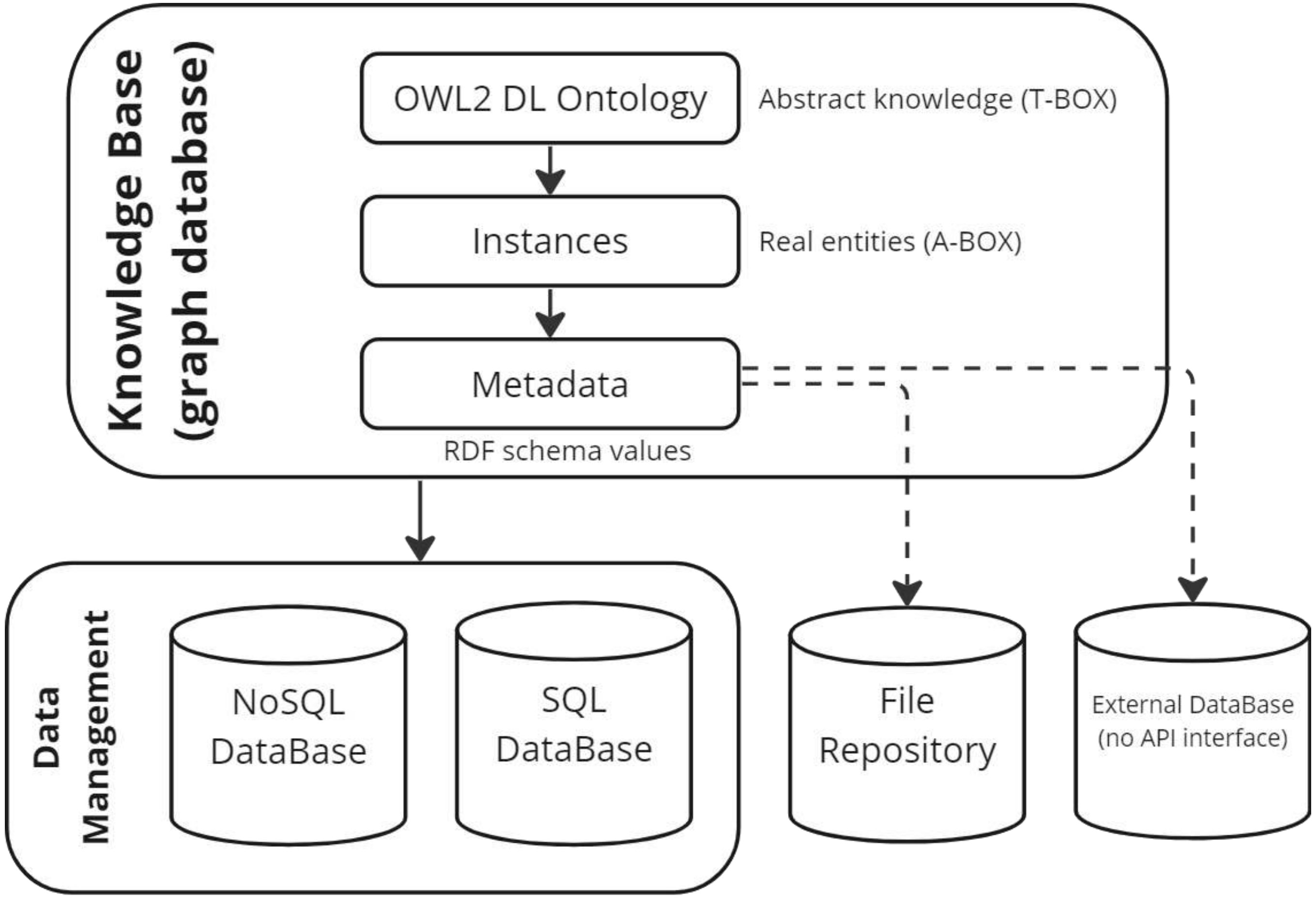
LOCAL DATA

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]).

1. Planning

o Did they plan a priory which data they needed, and how to extract knowledge from it?
o Did they work with Data Management Plans?

Co-BRAIN



Data Federation is possible i.e. data can be made **directly reachable** from the KB through mapping with the graph database vendor querying system.

Data are **not directly reachable** from the KB but provide users with information about their location and accessibility.

OTE Capabilities

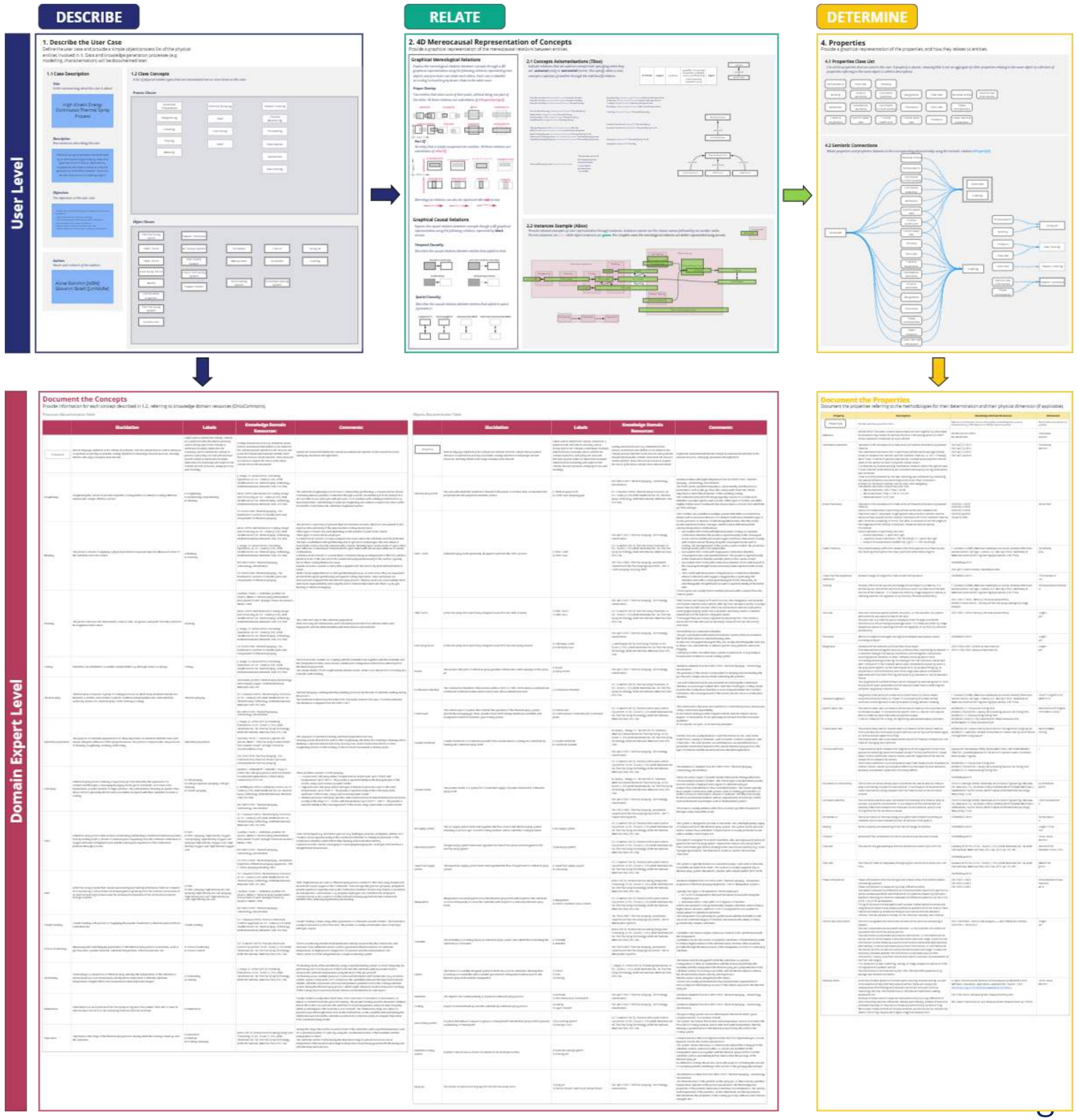
(from OntoTrans and EMMC)

1. **Representing** manufacturing process challenges in a standard ontological form as technical and business Innovation Cases
2. **Connecting** innovation cases with existing appropriate information sources i.e. available data and materials modelling solutions
3. **Recommending** consistent materials modelling workflow options
4. **Supporting** simulation and validation activities
5. **Providing** semantic results interpretation to facilitate sharing and re-use of innovation cases and results

1. Planning

o Did they look for existing semantic assets to prepare for interoperability?

How to **extract the knowledge from domain experts** for an ontologist to create a formal representation of the domain knowledge (**T-BOX** of a **Domain Ontology**)?



OntoTrans Conceptualisation Template!

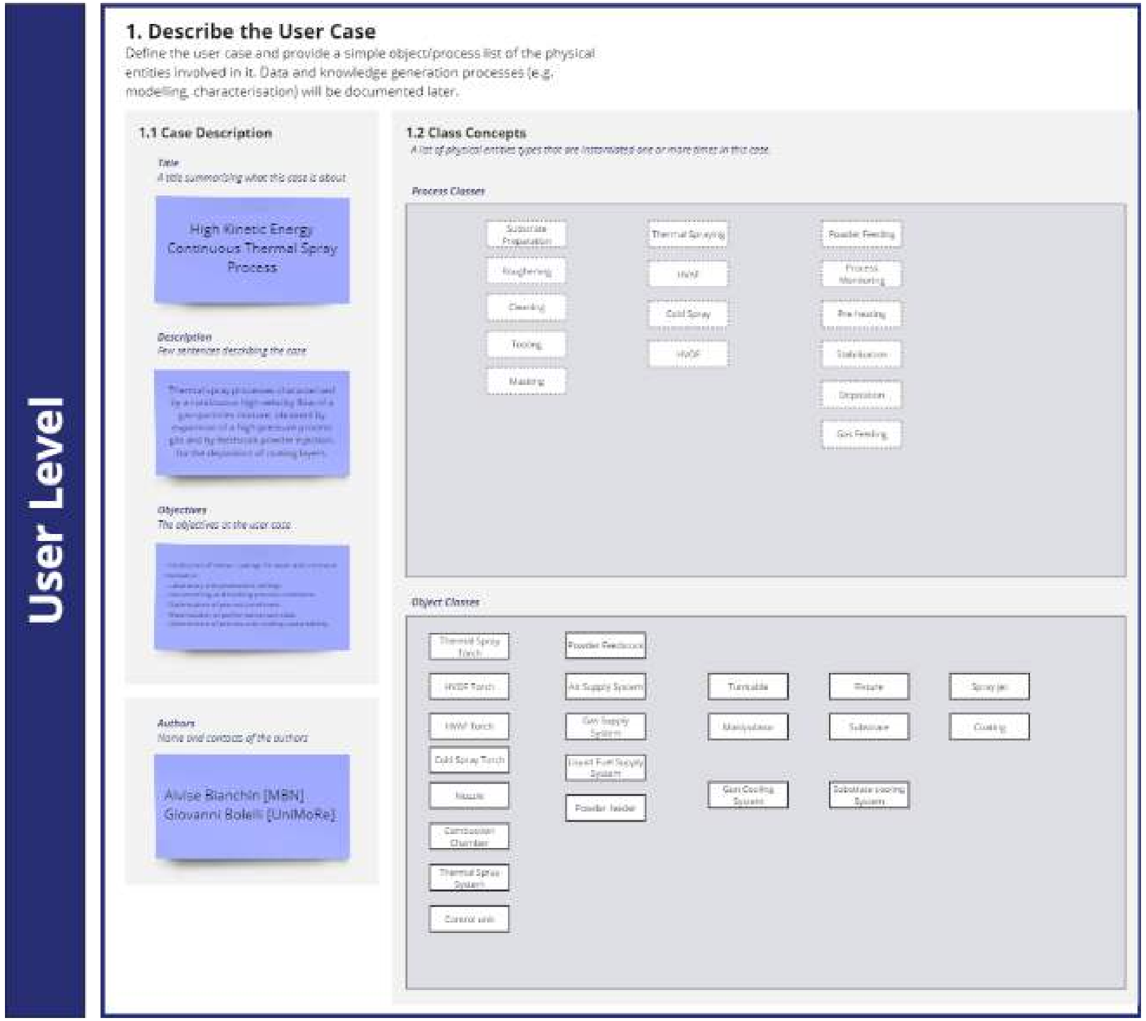
Gathering and formalising a domain knowledge is done through the process of **conceptualisation**, i.e., by identifying ontological concepts in the form of **classes**, **relations**, and **axiomatic constraints** that cover the domain of interest.

To overcome the **barriers** coming from the **lack of expertise** in ontology engineering, the **OntoTrans project** has developed a methodology for the interaction between the **translator** and the industrial stakeholders, aimed to facilitate collective contributions to the conceptualization effort.

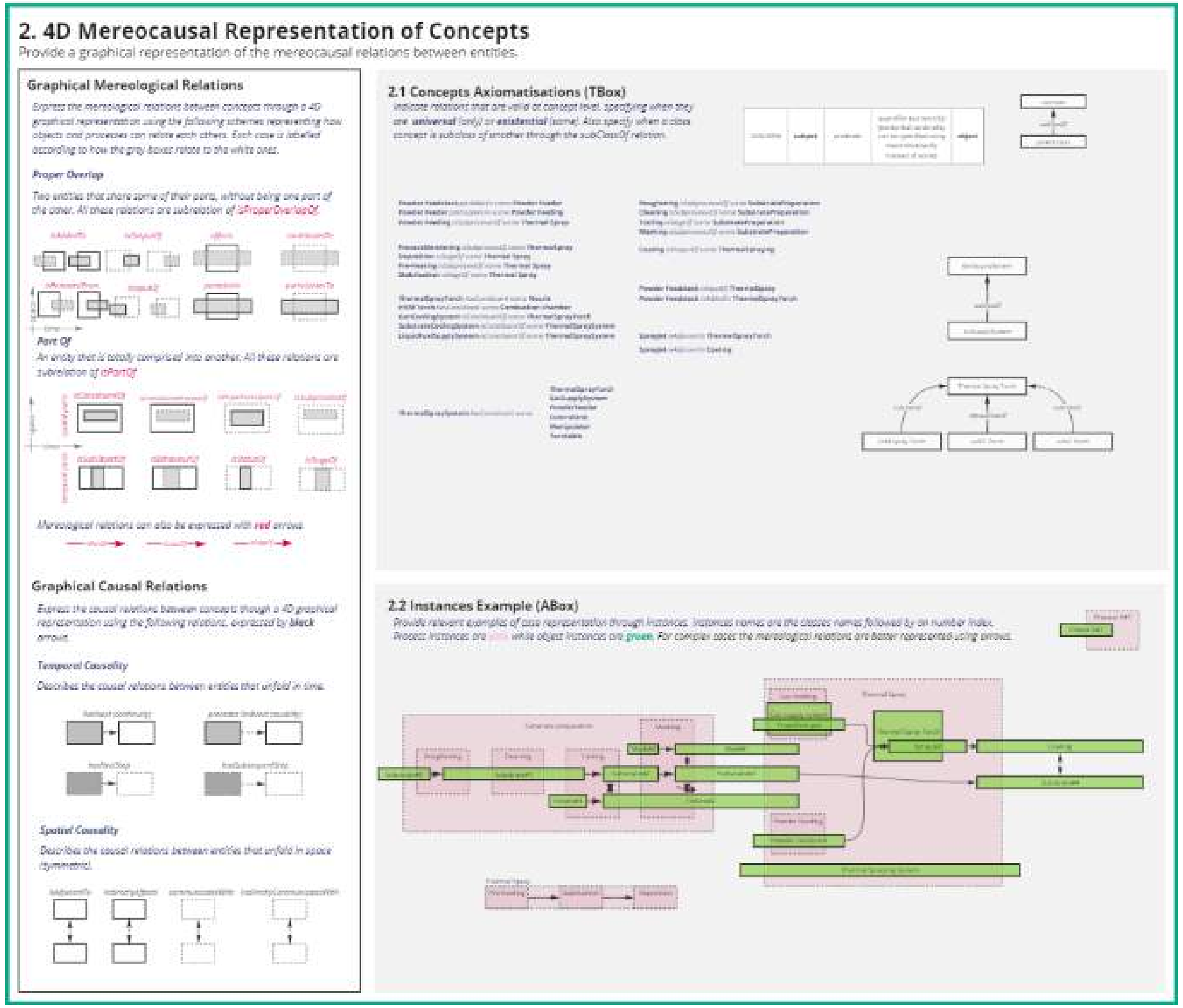
1. Planning

o Did they look for existing semantic assets to prepare for interoperability?

DESCRIBE



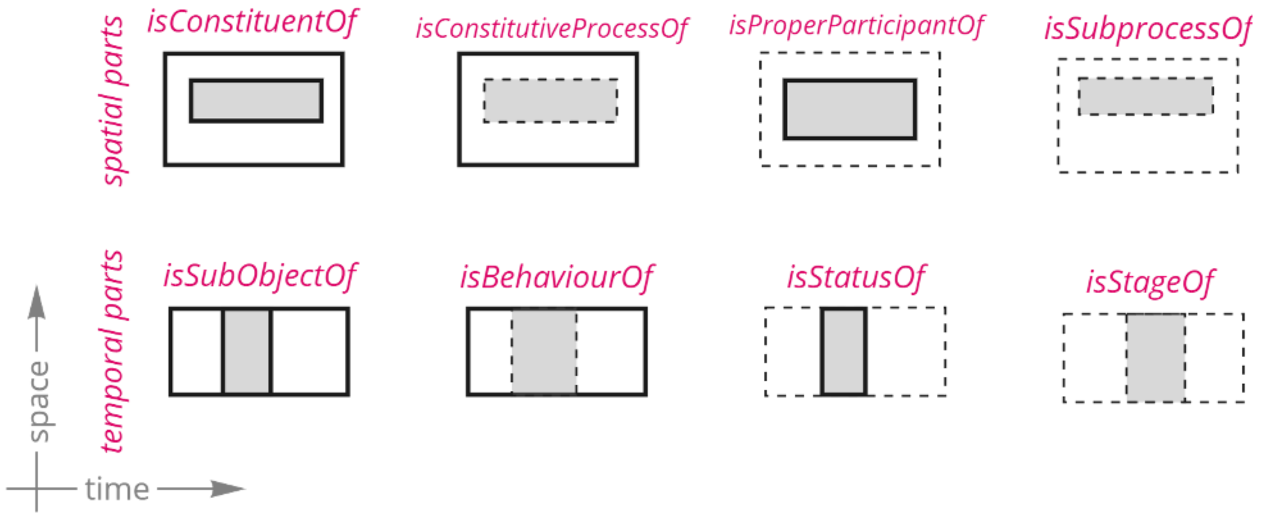
RELATE

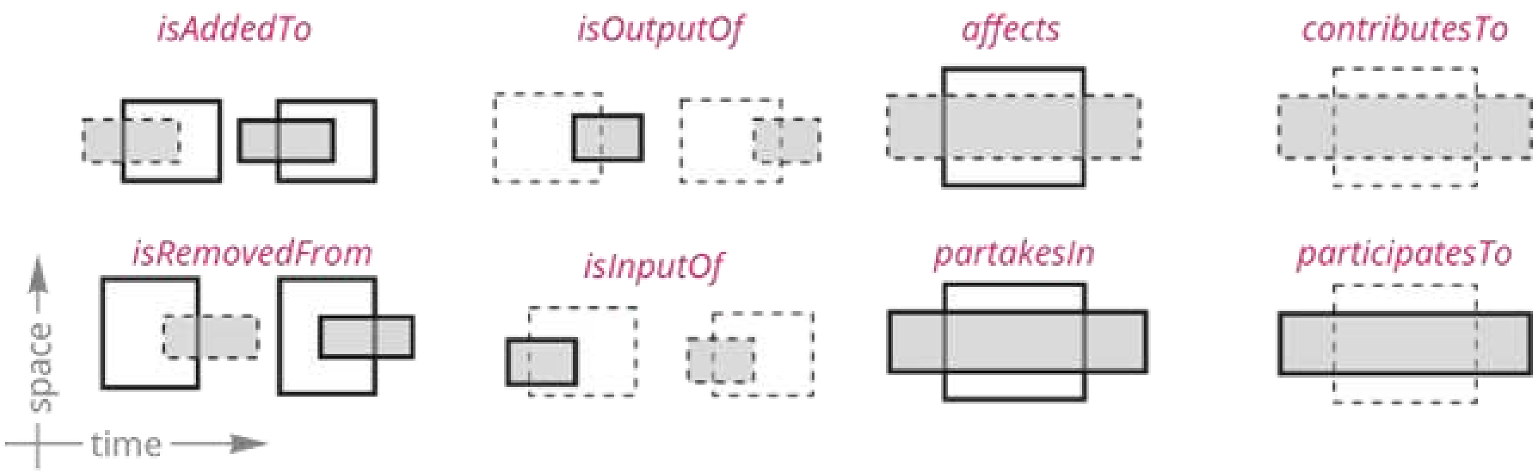


1. Planning

- Did they look for existing semantic assets to prepare for interoperability?

A simple **visual summary** of **EMMO fundamental ontological relations** has been provided to the users to let them understand how an ontology formalise the relations between entities.



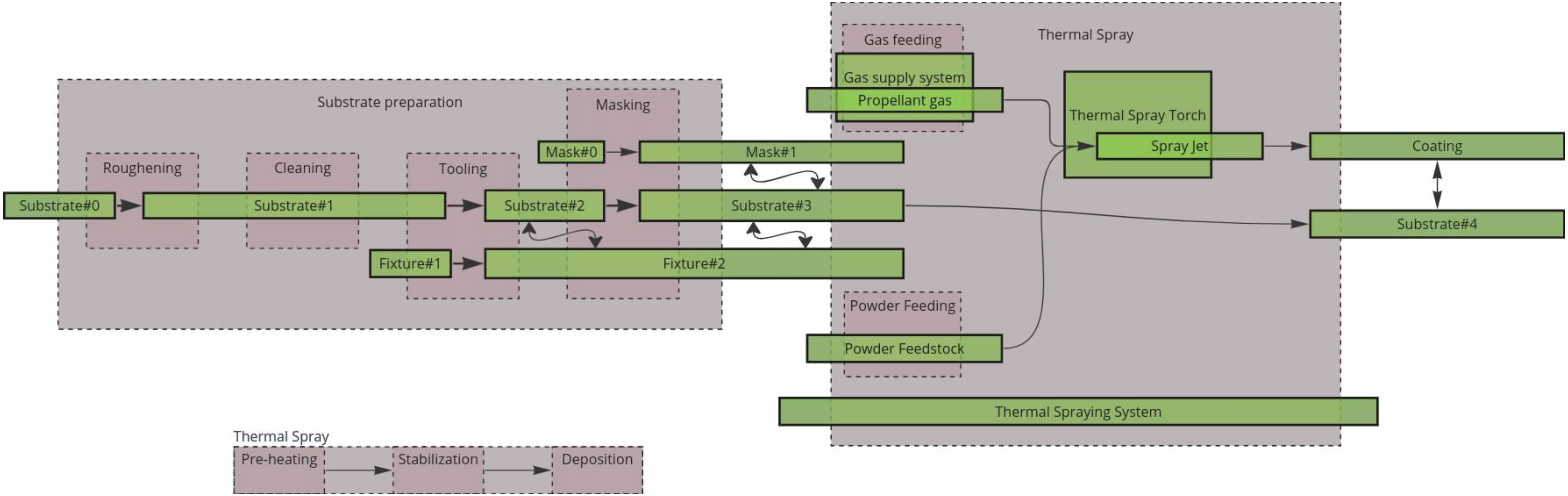


1. Planning

o Did they look for existing semantic assets to prepare for interoperability?

Co-BRAIN

4D conceptual representation of a **Thermal Spraying process**, connecting all the relevant entities using EMMO mereocausality relations, to document the **overall state of things**.

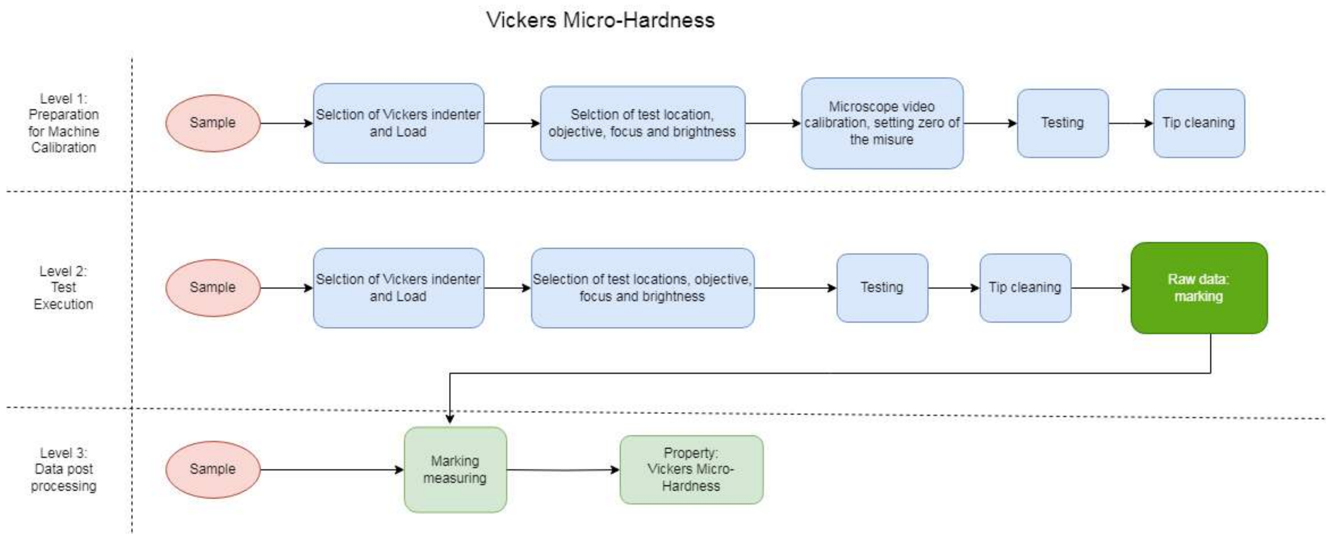


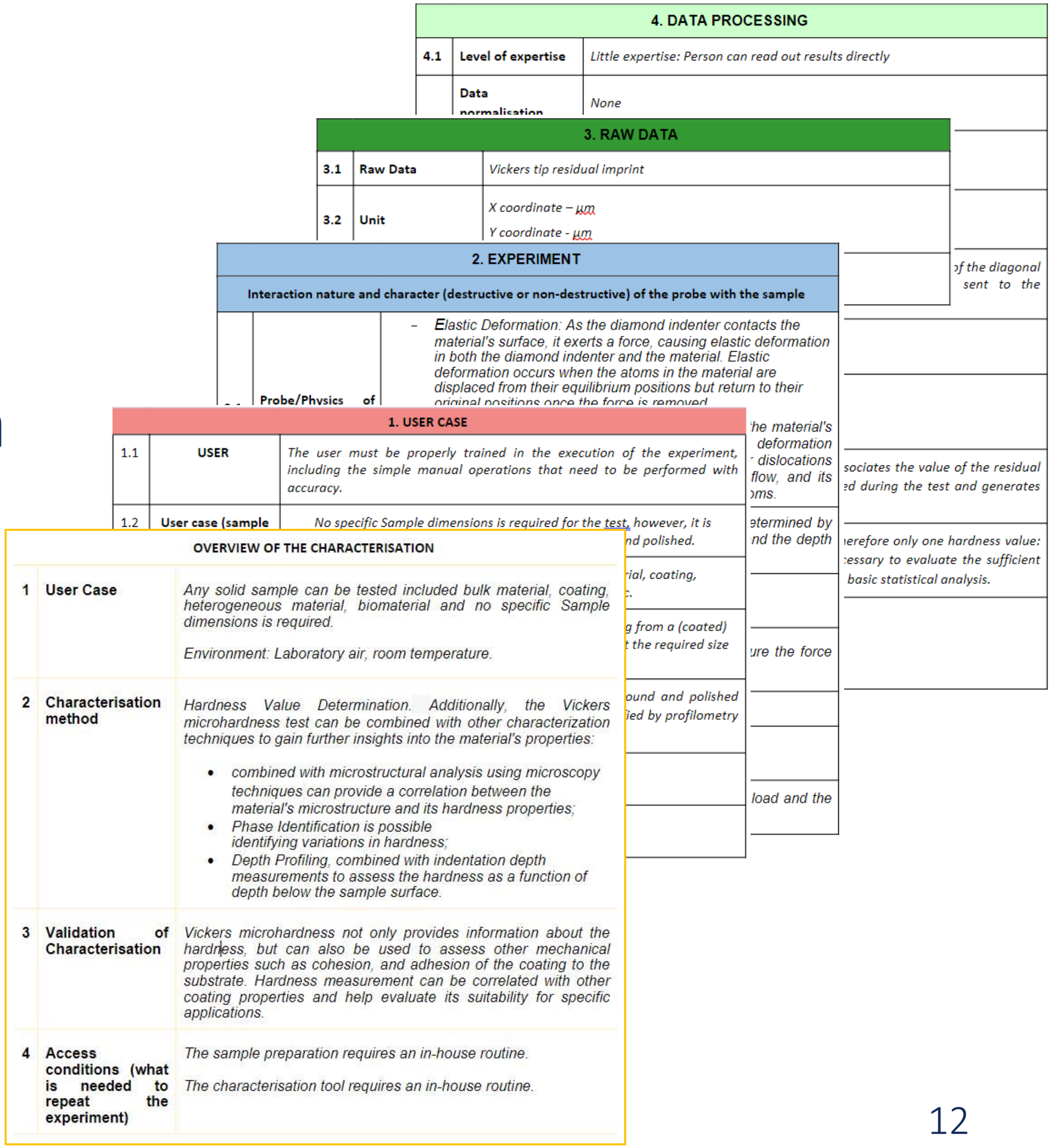
1. Planning

- Did they look for existing semantic assets to prepare for interoperability?

MODA and CHADA templates have been used mainly to **document modelling and characterisation workflows**.

Their ontological representation has been simplified since we had to focus more on data mapping and ingestion.



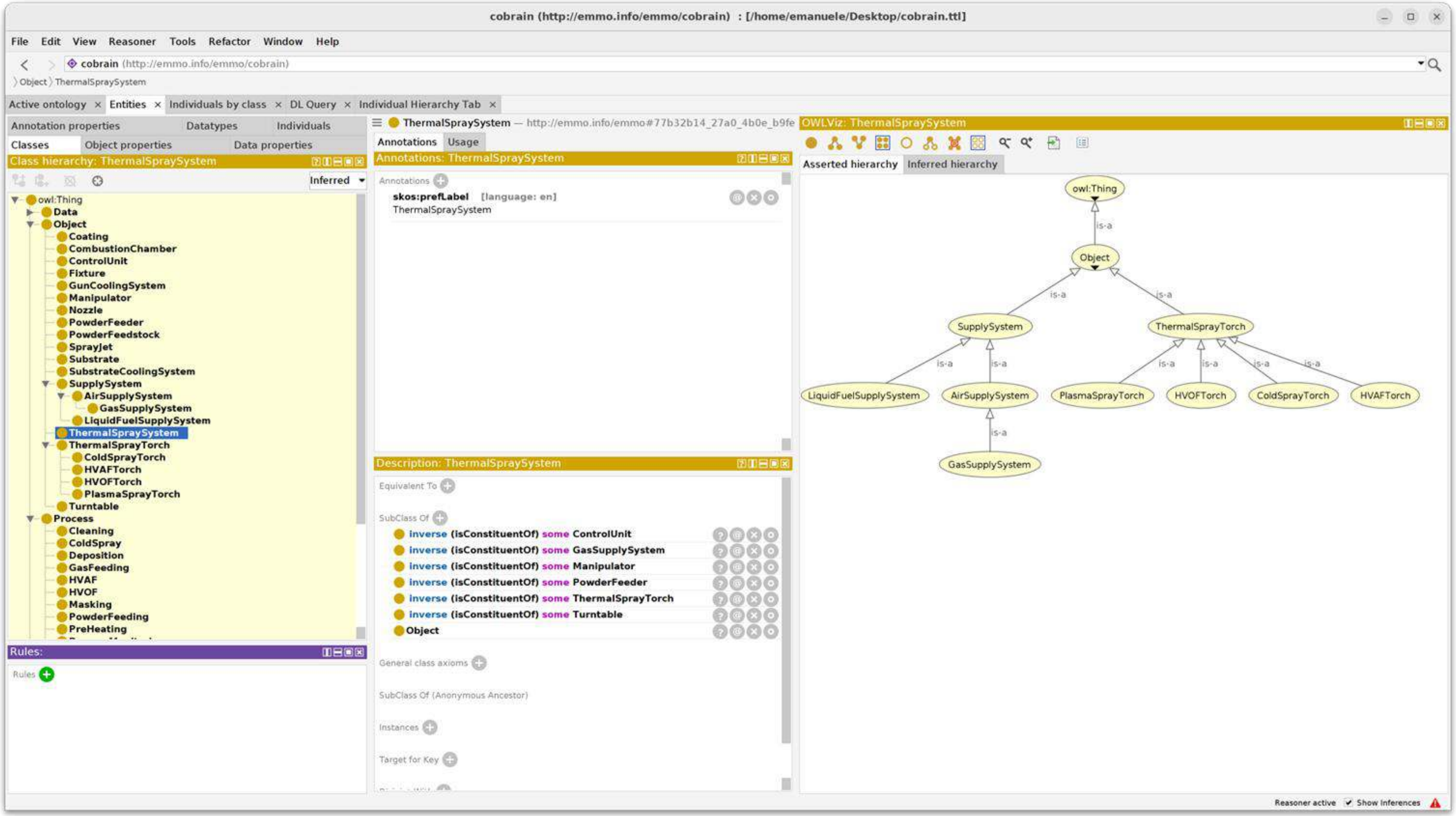


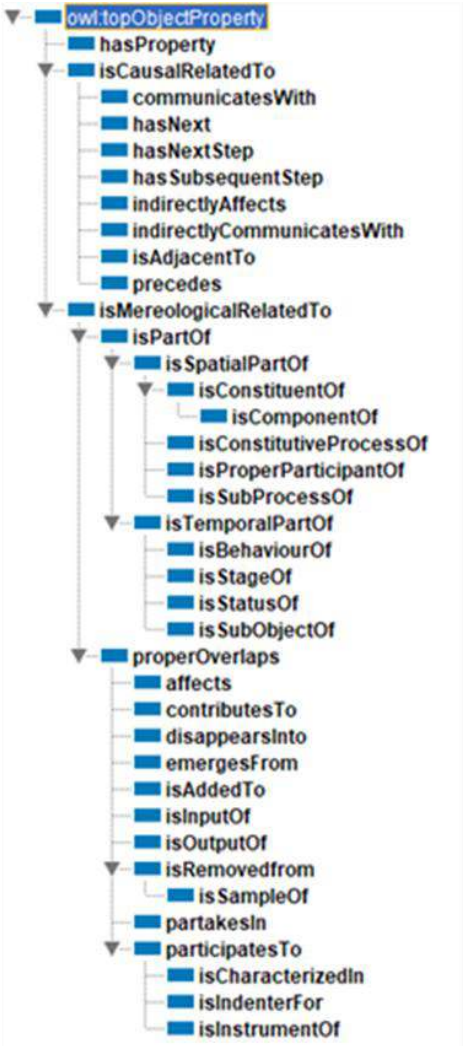
1. Planning

- Did they look for existing semantic assets to prepare for interoperability?

CoBRAIN Thermal Spraying Ontology

https://github.com/cobrain-project





1. Planning

o Did they look for existing semantic assets to prepare for interoperability?

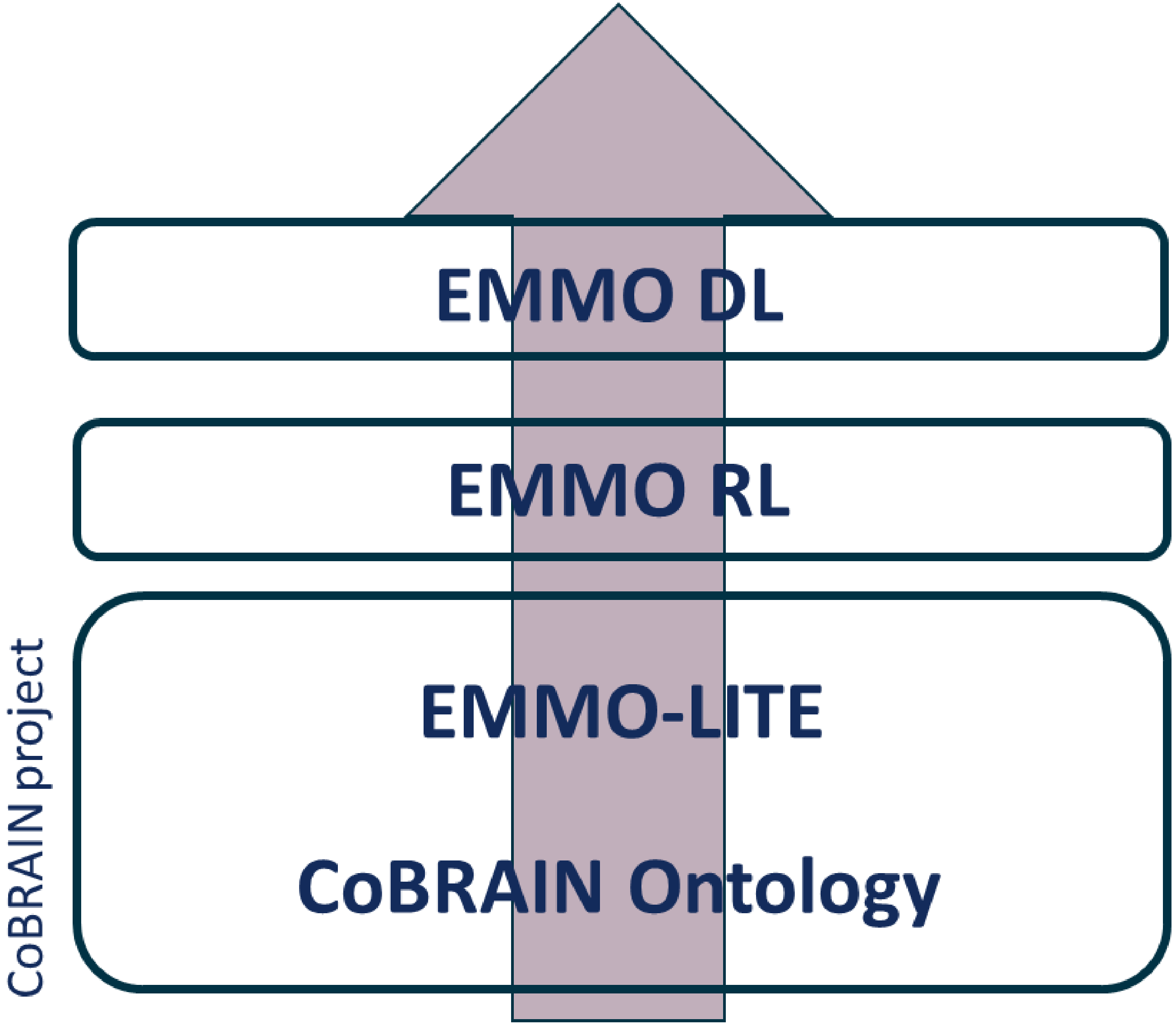
EMMO-LITE has been used as a reference ontology, to **facilitate usage by non-EMMO experts**

EMMO-LITE Public

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

ontology Public

https://github.com/cobrain-project/ontology



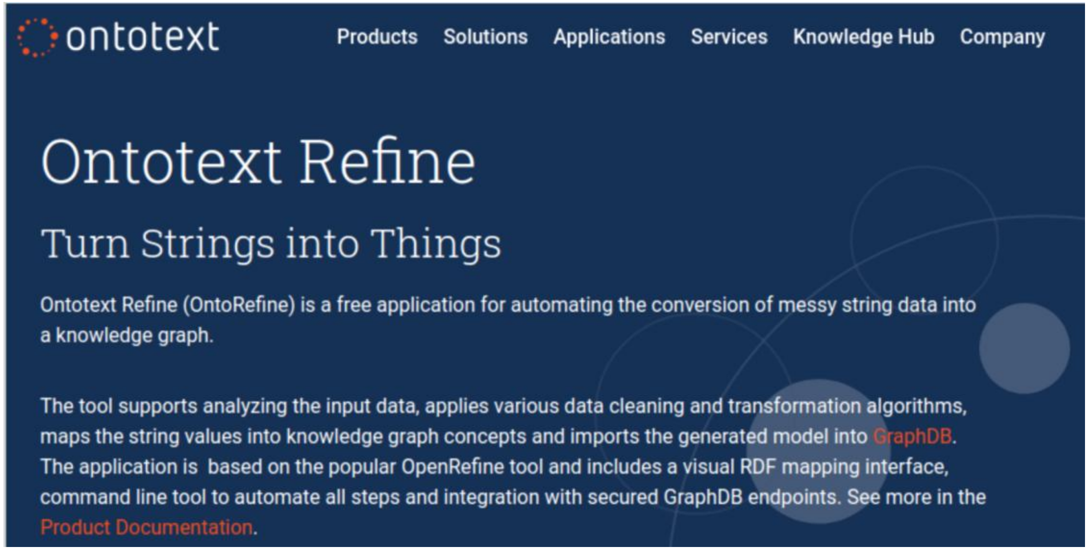
1. Planning

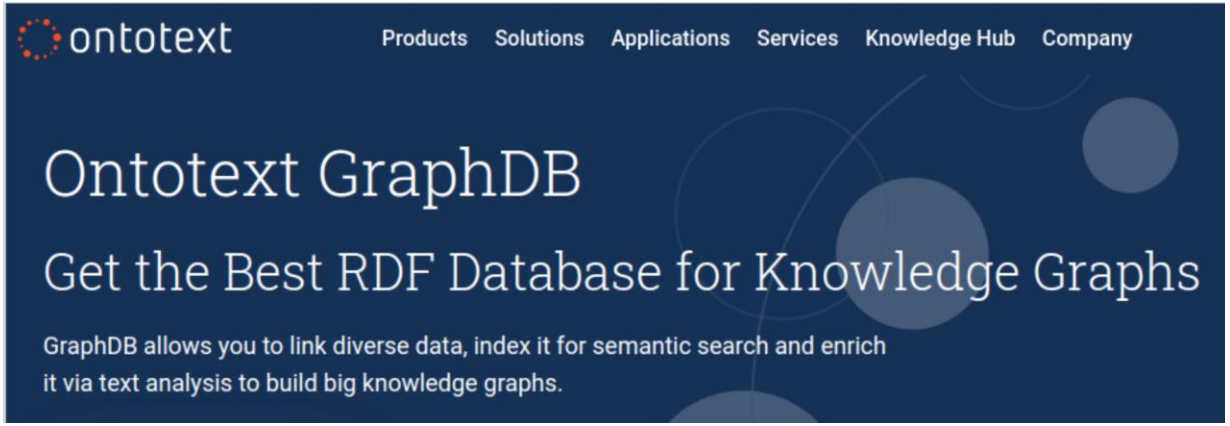
- Did they purchase software and or equipment to produce data?

Key point for the CoBRAIN **Knowledge Management Platform** design:

No wheel reinventing!

The overall system must rely on existing commercial tools with free-to-use licensing option and W3C standards that have reached a significant level of maturity.

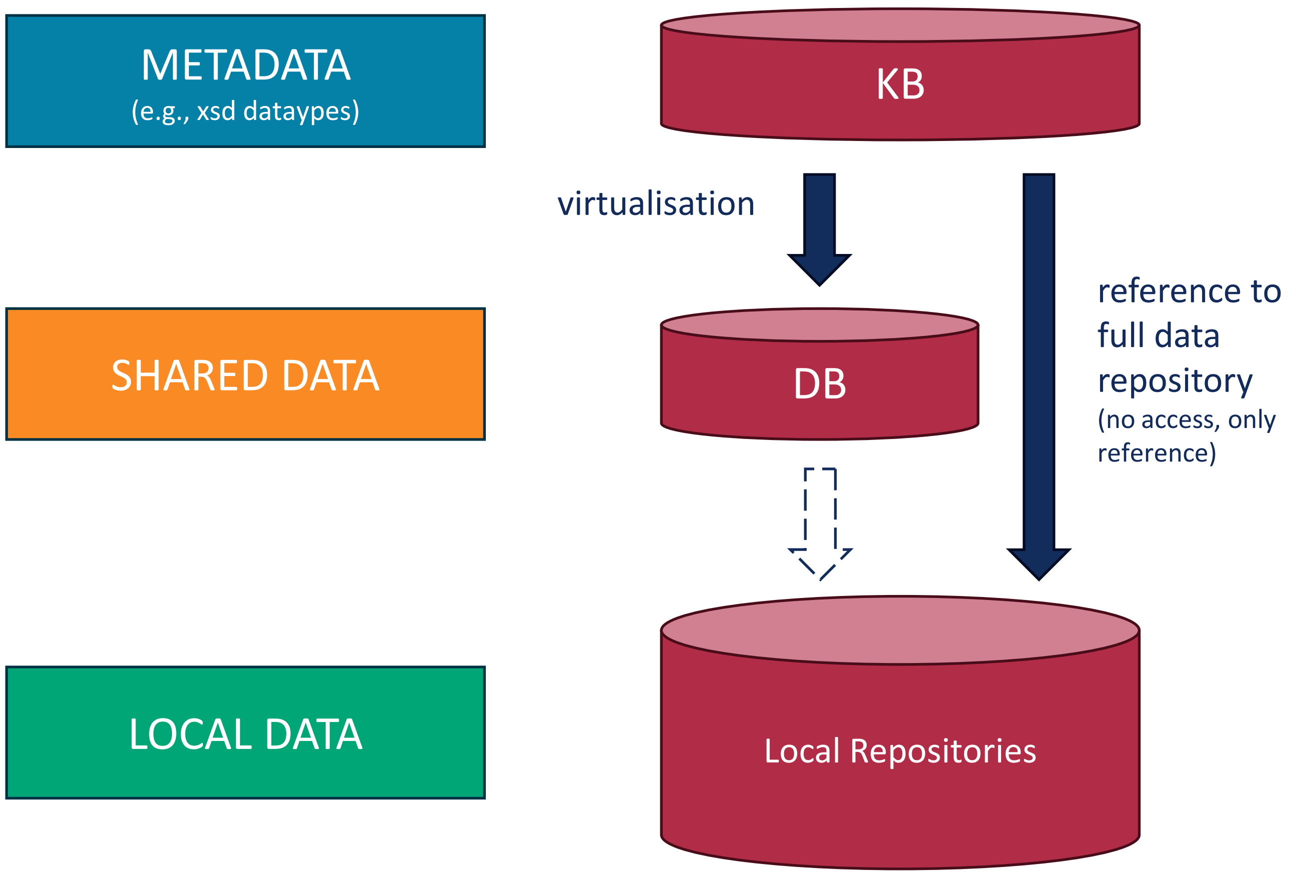




2. Acquiring

- How is data acquired?
- Are meta data acquired?

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. Acquiring

- How is data acquired?
- Are meta data acquired?

Co-BRAIN

METADATA
(stored in KB as e.g., xsd dataypes)

SHARED DATA
(stored in KB as e.g., xsd dataypes)

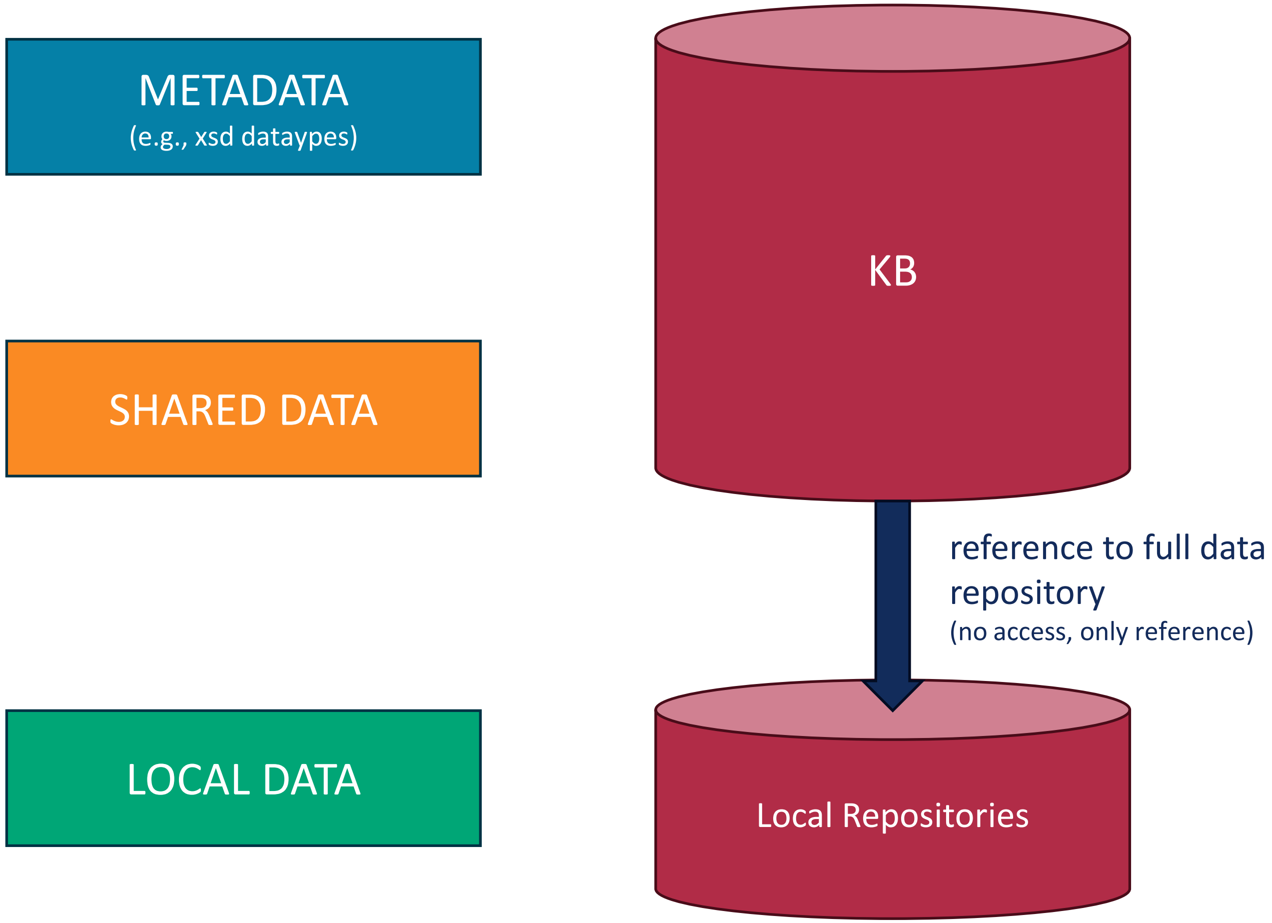
SHARED DATA and METADATA

Field	Description
Measurement type	**String**: “Volume di una buca o un picco:” followed by a shortened form of the string indicating the name of the original file and the elaboration operation
File name	**String**: Original profilometer file including path
Date/time	**yyyy/mm/dd hh:mm:ss**: date and time of elaboration
Hole surface	**μm^2**: area of the projected surface occupied by peaks, followed by a string listing the complete sequence of operations leading to the elaborated profile
Peak surface	**μm^2**: area of the projected surface occupied by peaks, followed by a string listing the complete sequence of operations leading to the elaborated profile
Hole volume	**μm^3**: overall volume of the hole (below the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Peak volume	**μm^3**: overall volume of the peaks (above the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Maximum depth (hole)	**µm**: maximum depth of the hole (below the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Maximum height (peak)	**µm**: maximum height of the hole (below the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Average depth (hole)	**µm**: average depth of the hole (below the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Average height (peak)	**µm**: average height of the hole (below the reference plane), followed by a string listing the complete sequence of operations leading to the elaborated profile
Objective	**String**: Magnification of the objective employed in the measurement [optional, user-typed, inserted in the first measurement file]
Step	**µm**: z-scan step size [optional, user-typed, inserted in the first measurement file]

2. Acquiring

- How is data acquired?
- Are meta data acquired?

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]).

2. Acquiring

- How is data acquired?
- Are meta data acquired?

Co-BRAIN

Data are acquired and **stored locally** using **partner's specific methods and formats**.

Most of them are **Excel tables**. Some data are in **JSON format** (unfortunately not JSON-LD).

Example of tabular data providing HVOF process details.

Sample ID	Nominal Composition	Lot. N.	Substrate	Date	Run n°	length mm	width mm	FR H_2 FMR	p H_2 psi	FR O_2 FMR	p O2 psi	FR [illegible] FM	p Air psi	Pwdr. Feed g/min	Standoff mm	Pitch mm	Speed mm/s	n° cycles	passes/cycle	passes	weight in g	weight fin g	weight in g	weight fin g	weight in g	weight fin g	gain g	sprayed mass g	efficiency g
Al0(CrMnFeNi)+60TiC RUN1	Al0(Cr20Mn25Fe40Ni15)86+60vol%TiC	1364-N-23	AISI 304	19/02/2024	1	60	25	56	140	30	170	35	100	20	250	5	750	13	3	3	34,67	37,76	34,35	37,5	34,26	37,35	3,11	5,2	59,8

2. Acquiring

- How is data acquired?
- Are meta data acquired?

Co-BRAIN

Ideal solution: a new project that implements **a full framework for assisted/automated data collection** for all CoBRAIN datasets, providing a built-in semantic enhancement of data

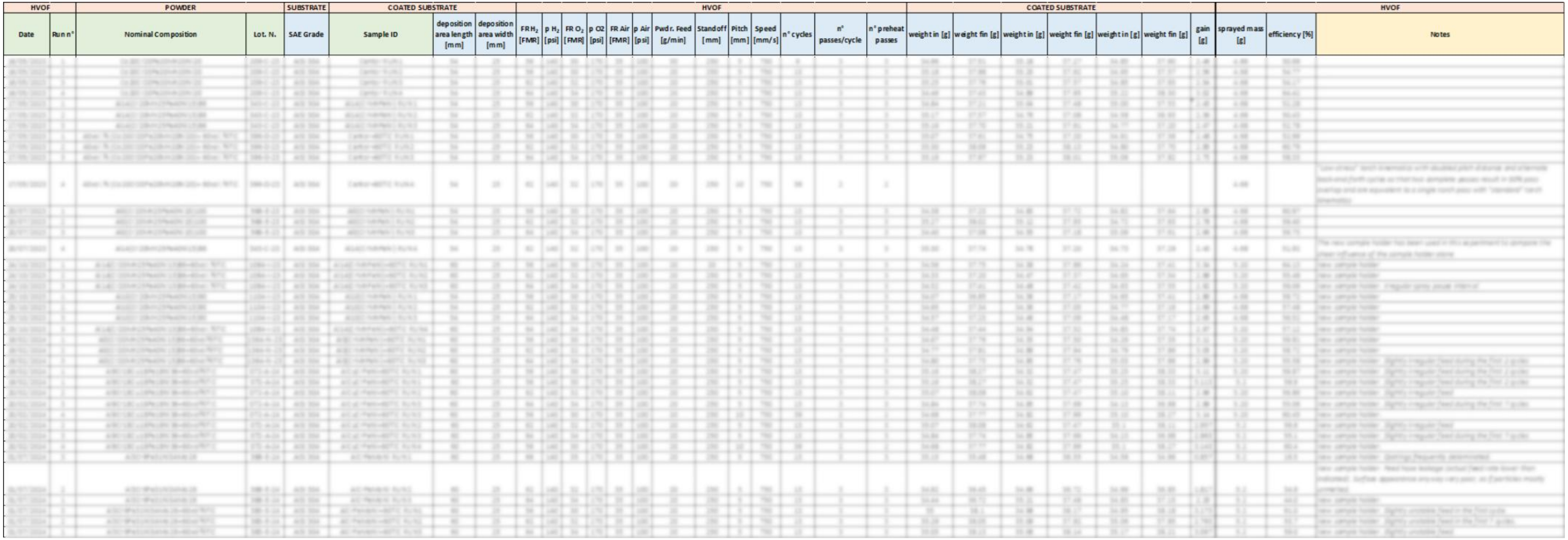
Proposed Solution: UNIBO will **refactor all collected data**, providing as-much-as-possible **constrained templates** for data collection ("if you can't beat them, join them")

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
FCC0q8	[[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	XIUQ32	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	XIUQ32	W	[[1,3,4],[2,1,5]]	5	3

2. Acquiring

- What types of data are there?
- Are the data described as and when they come in?

Excel files have been reviewed and templates for the acquisition of data for each dataset type have been provided by UNIBO after meetings with partners (still WIP)



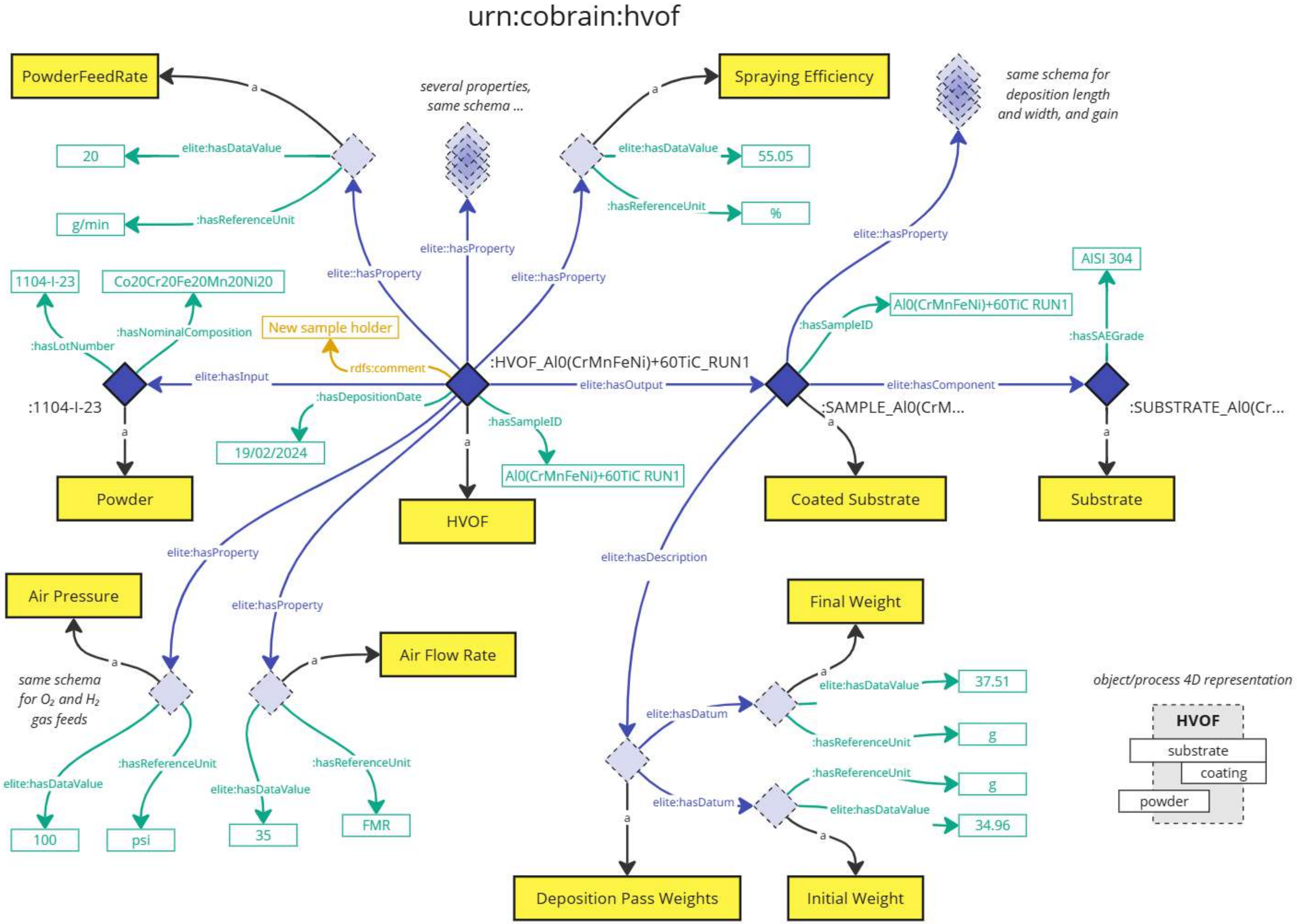
3. Processing

- How are data extracted?
- Are there automatic workflows in place?
- How is provenance ensured?

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



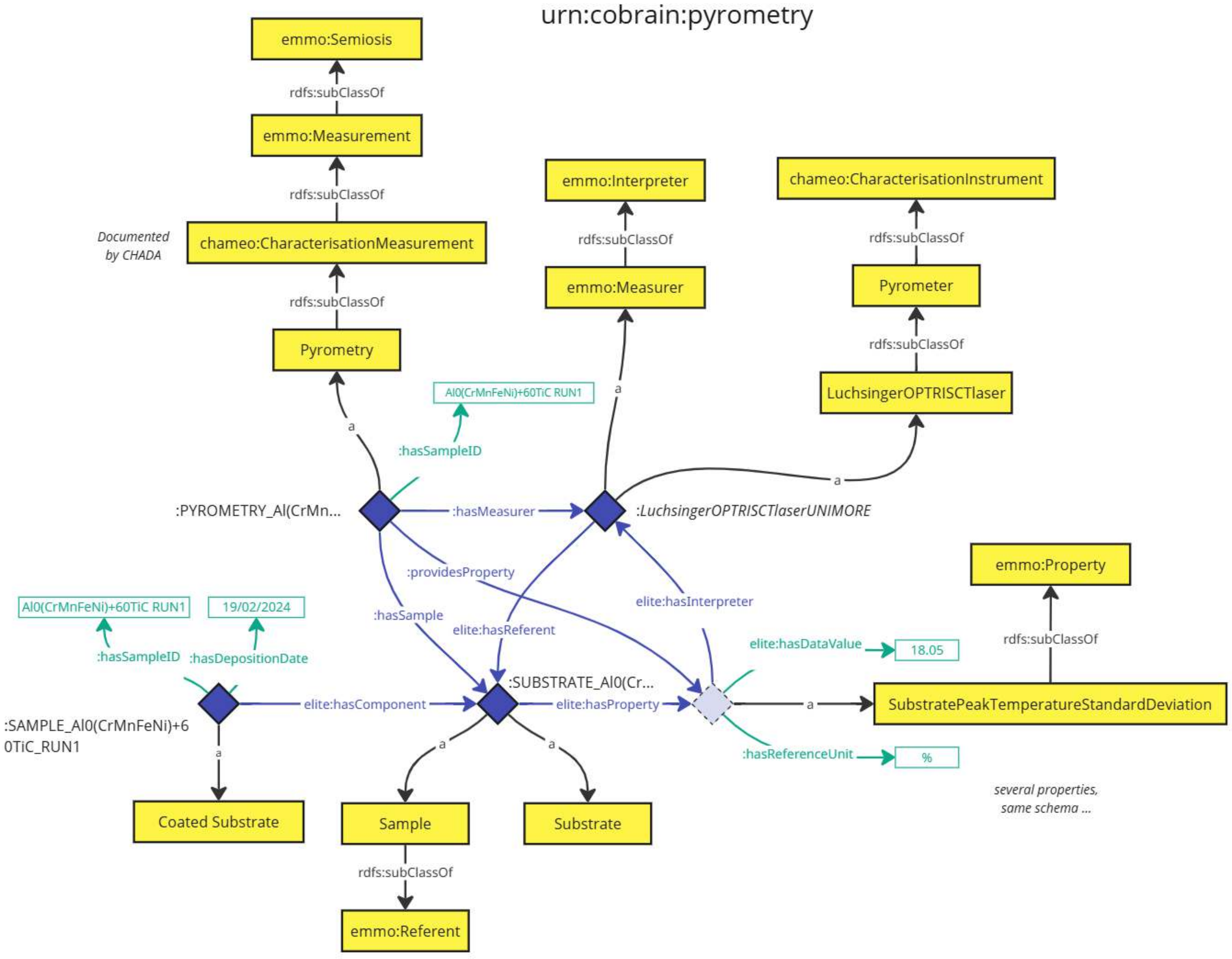
3. Processing

- How are data extracted?
- Are there automatic workflows in place?
- How is provenance ensured?

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

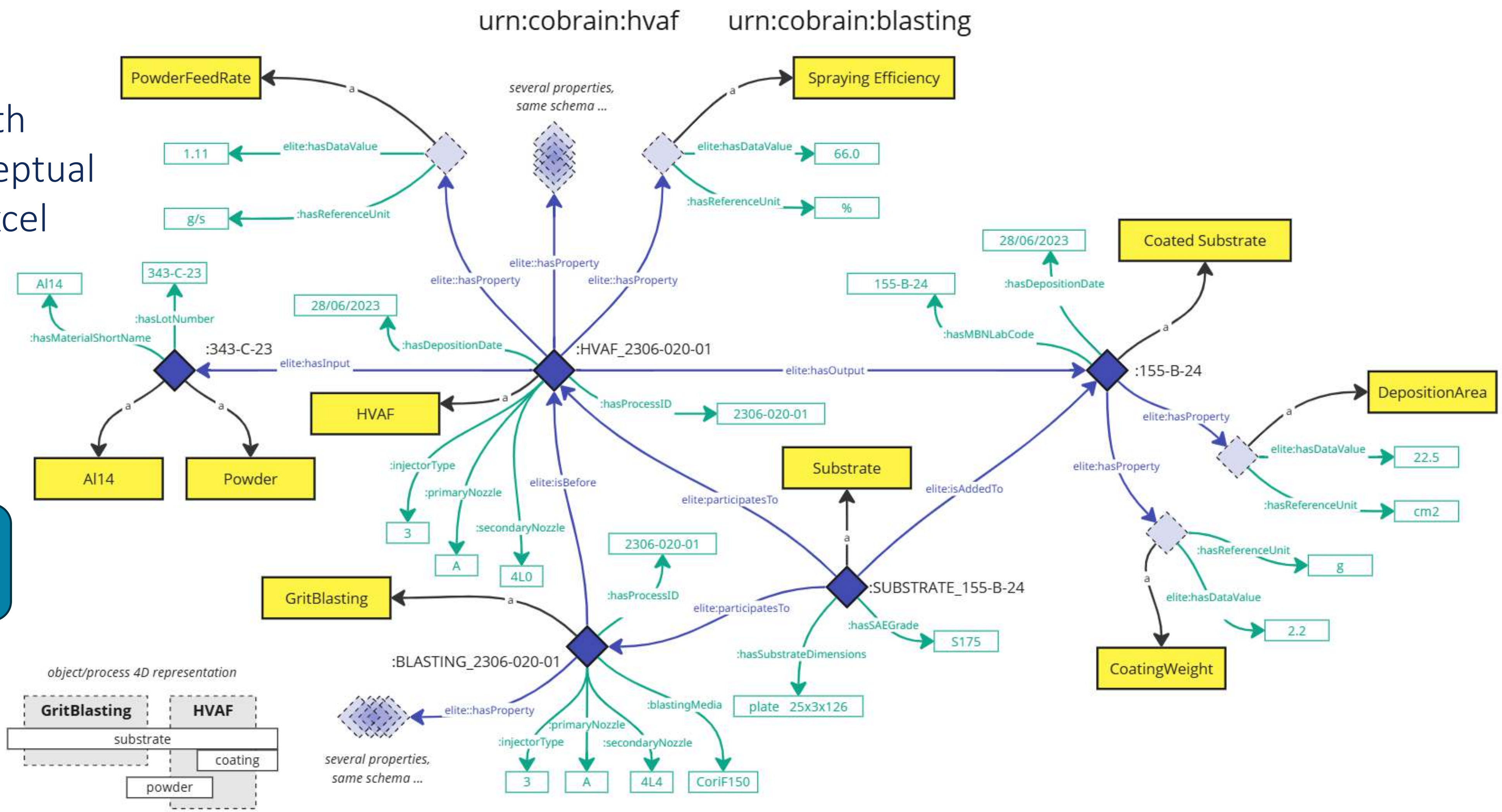
- How are data extracted?
- Are there automatic workflows in place?
- How is provenance ensured?

Co-BRAIN

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.

HVAF Thermal Spraying Logbook + Grit Blasting



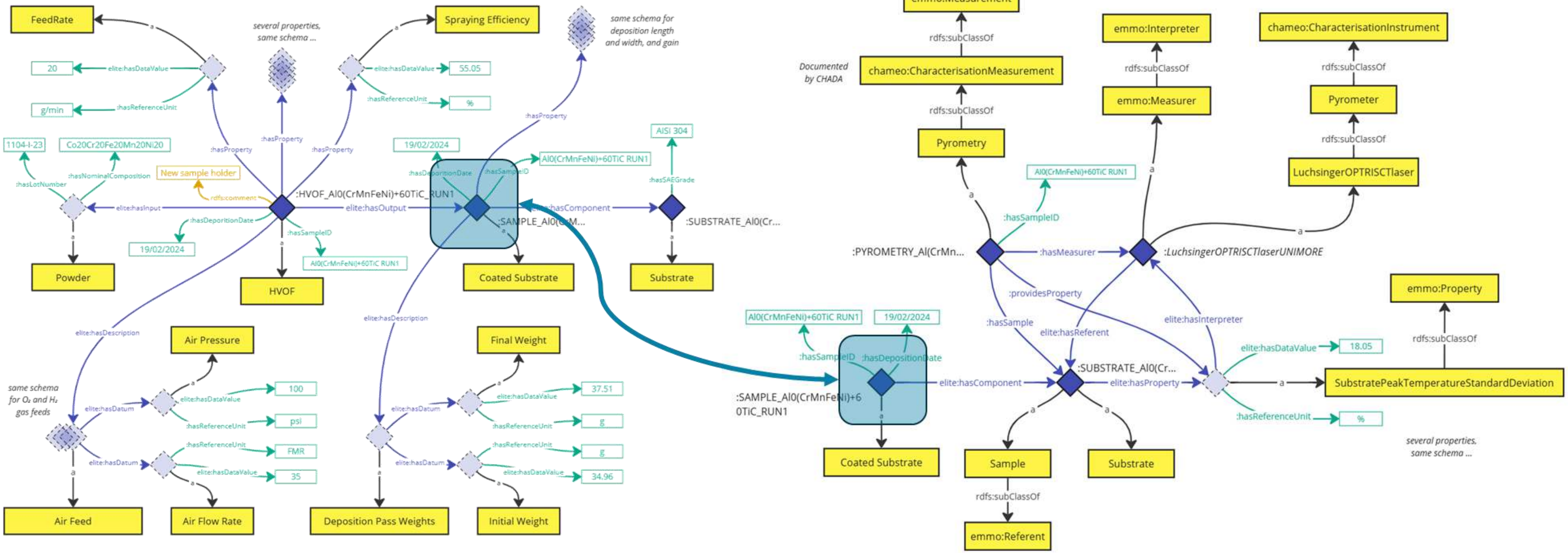
3. Processing

- How are data extracted?
- Are there automatic workflows in place?
- How is provenance ensured?

Co-BRAIN

UMR HVOF Thermal Spraying Logbook

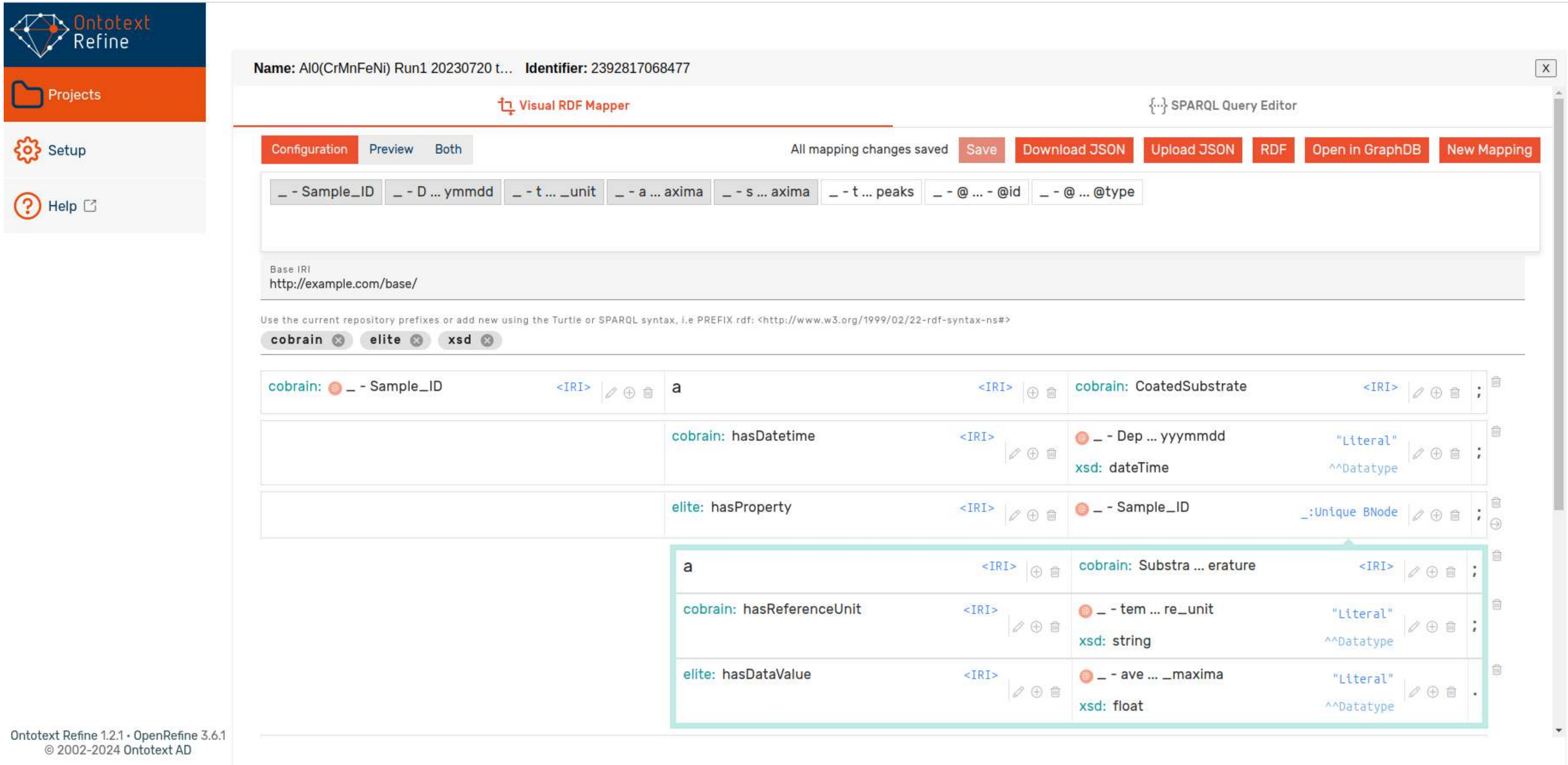
Optical pyrometer Luchsinger OPTRIS CTlaser



3. Processing

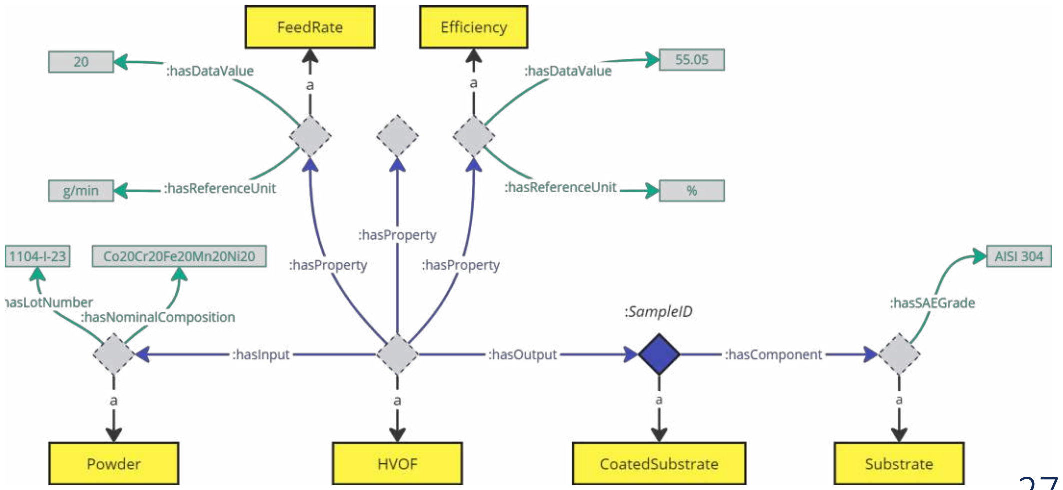
- How are data extracted?
- Are there automatic workflows in place?
- How is provenance ensured?

Mapping of HVOF Pyrometry UNIMORE T3.2 with OntoRefine



We have found some practical issues in the excel-oriented approach experimental and characterisation datasets:

- 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



SPARQL stands for SPARQL Protocol and RDF Query Language

SPARQL is a standard semantic query language for databases used for manipulating and retrieving data in the RDF (Resource Description Framework) format

RDF is a data model for representing information as a set of triples.

A triple consists of a subject, predicate, and object:

- **Subject**: Represents a resource (e.g., a person, a place).
- **Predicate**: Describes a property or relationship of the subject.
- **Object**: The value of the property or the resource related to the subject.

Subjects and predicates are always URI identifiers, but objects can be URIs or literal values.

4. Analysing

What tools did/will your project develop to extract knowledge from your data?

Co-BRAIN

```
PREFIX mov: <http://example.org/ontology/movies#>

SELECT DISTINCT ?movie ?genre WHERE {

    ?movie mov:director mov:StevenSpielberg .
    ?movie mov:genre ?genre .
    ?movie mov:releaseYear ?releaseYear .

    FILTER (?releaseYear > 2000) .

}
```

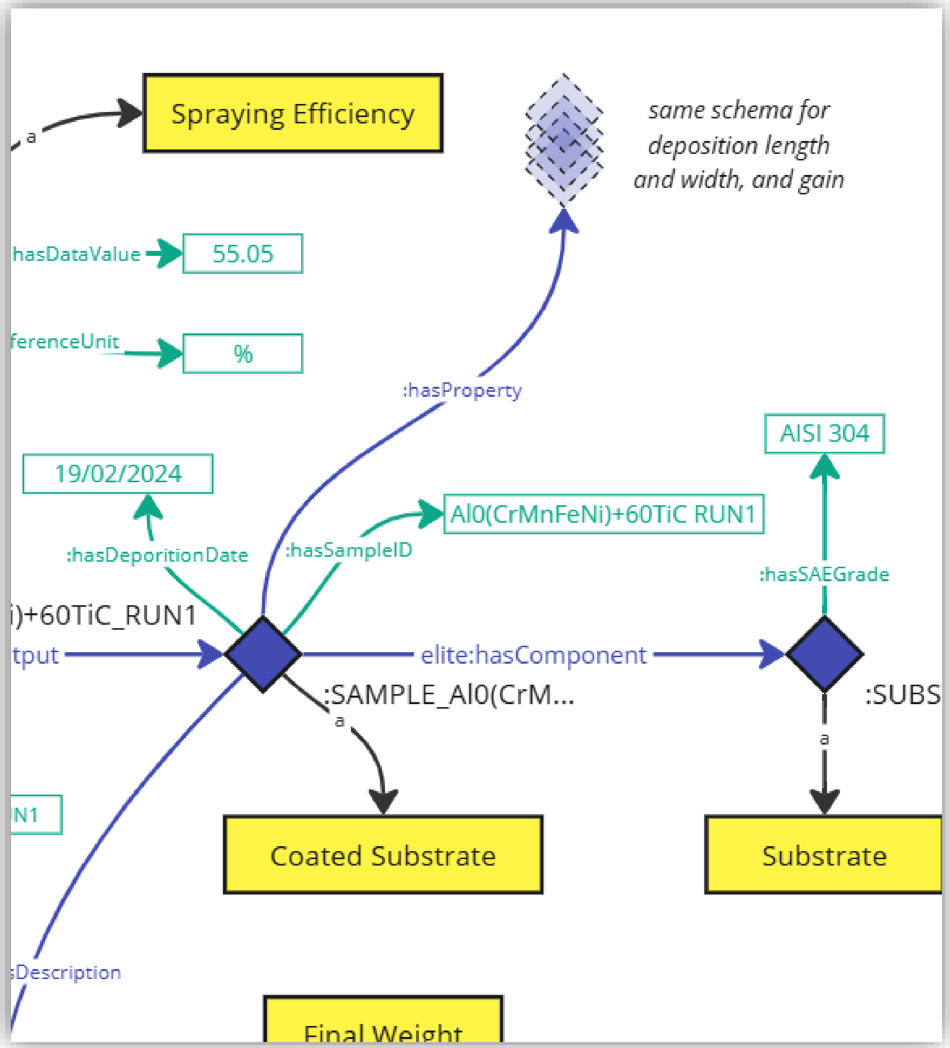
- Prefixes allow to substitute the full IRI with a concise version
- Select the desired elements, in this case without duplicates
- Predicates used to retrieve information from the graph
- Filtering clause that applies a condition to retrieve specific results

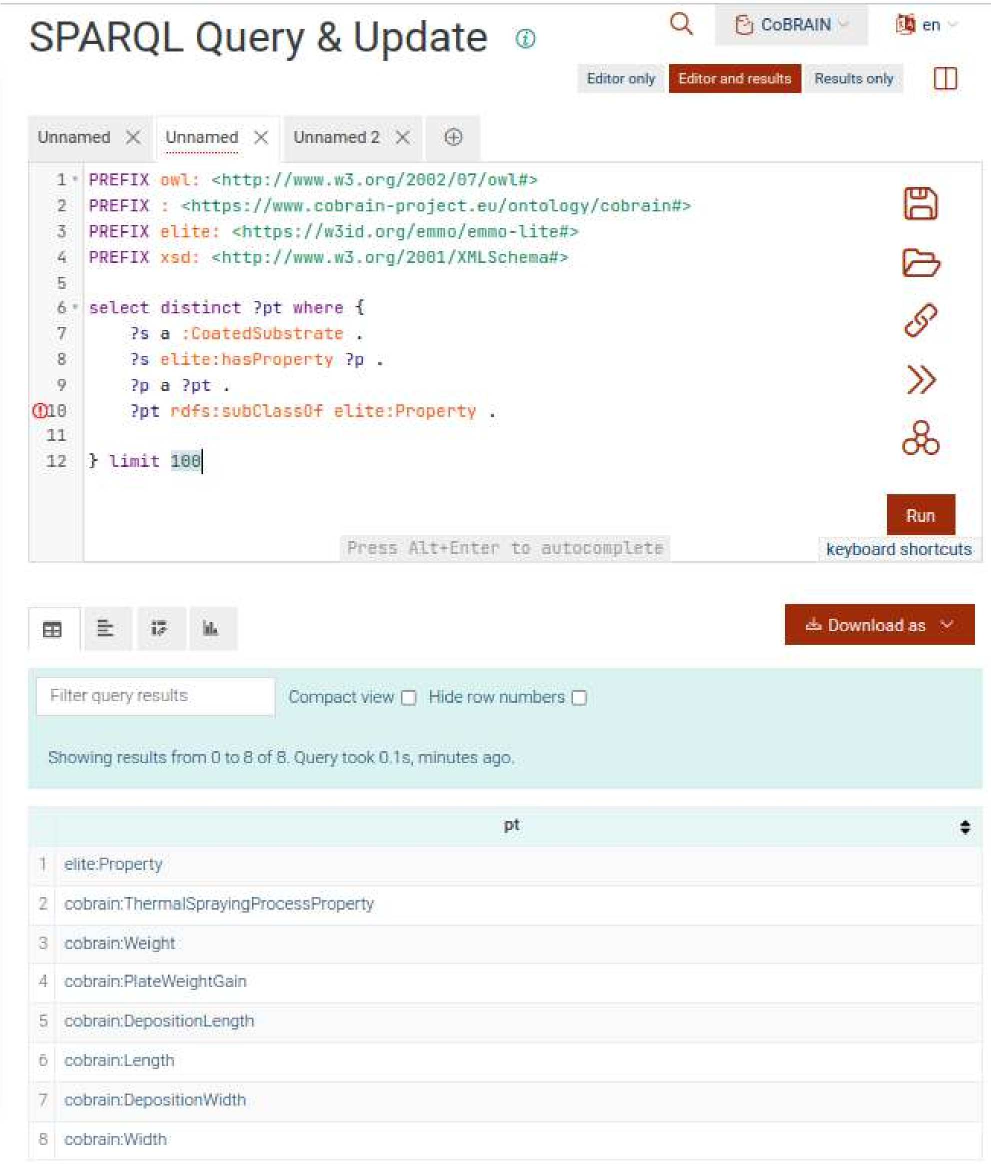
4. Analysing

What tools did/will your project develop to extract knowledge from your data?

Co-BRAIN

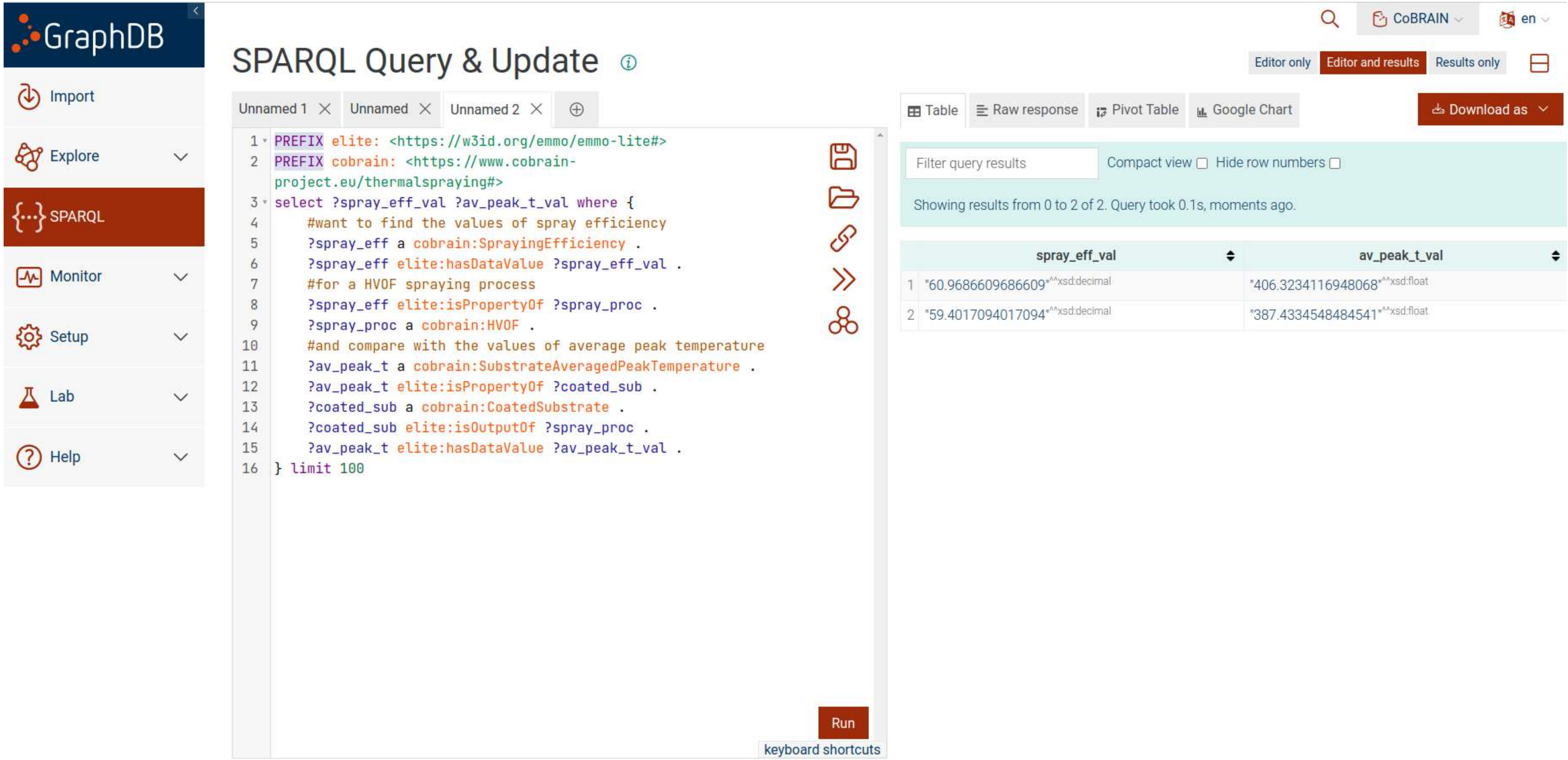
Conceptual schemes are paramount to elucidate the **CoBRAIN ontology data model** and **concept relations**, in order to **enable non-expert users** to generate the **correct SPARQL query** for their needs.





Co-BRAIN

- SPARQL Query Example



4. Analysing

What tools did/will your project develop to extract knowledge from your data?

Co-BRAIN

CoBRAIN Knowledge Base will be used mainly to:

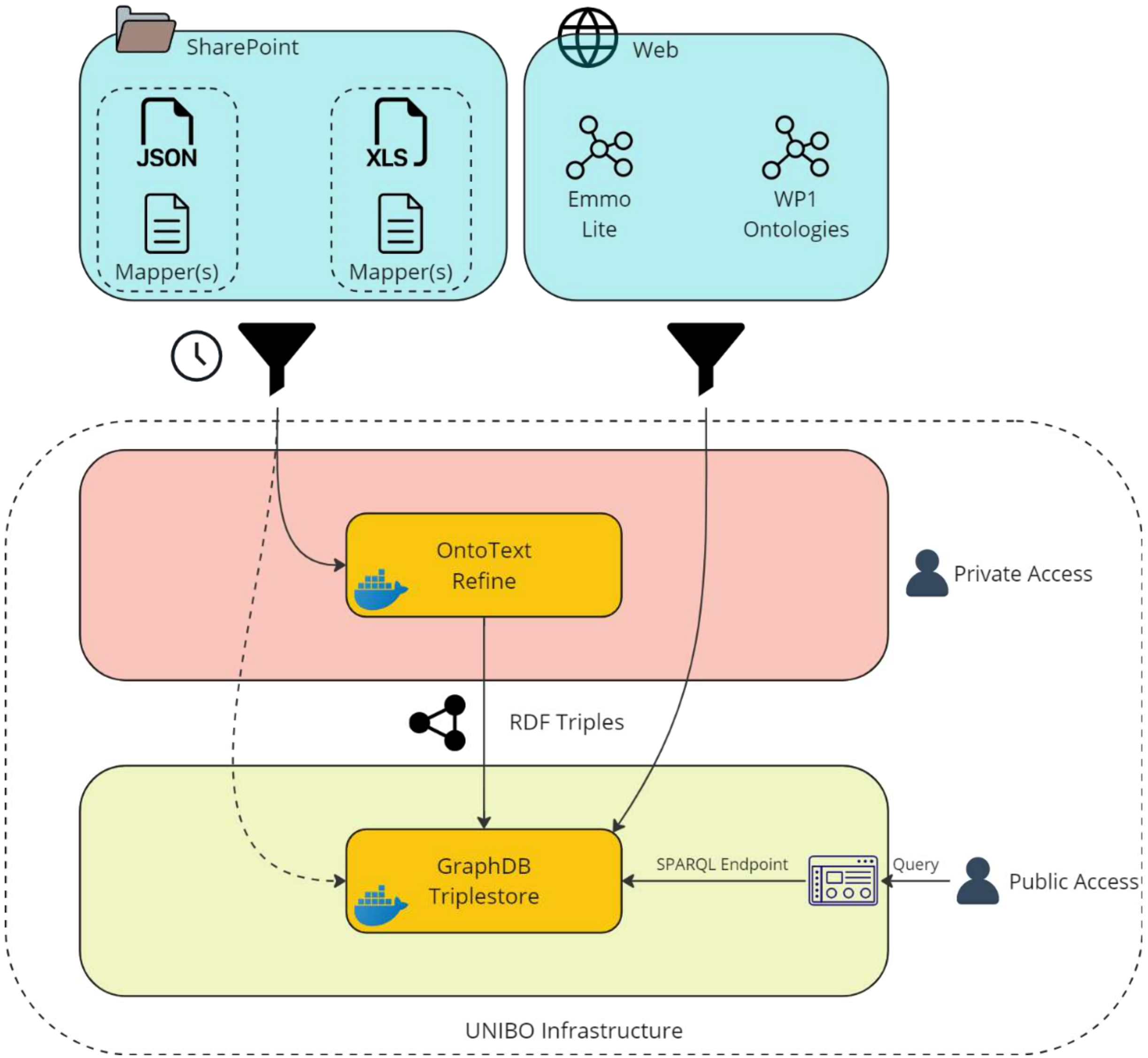
- Create a **federated interoperable database** for the CoBRAIN partners that can be used also after the project ends
- Provide **FAIR solution** for the dissemination and exploitation of project data
- Support the creation of a **SDSS** (Sustainable Decision Support System)

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

5. Preserving

- o How do you document your "new" data?
- o Are you using versioning?
- o In what sort of repository does the preservation happen?

Co-BRAIN



Ingestion Point

- Periodic import of experimental data and mappings from the project sharepoint
- Import of EMMO and project ontologies from websites

OntoTextRefine

- Conversion of raw data into formatted RDF triples
- Automatic saving on triplestore

GraphDB

- Storing of concepts and individuals with data
- Dashboard for query submission

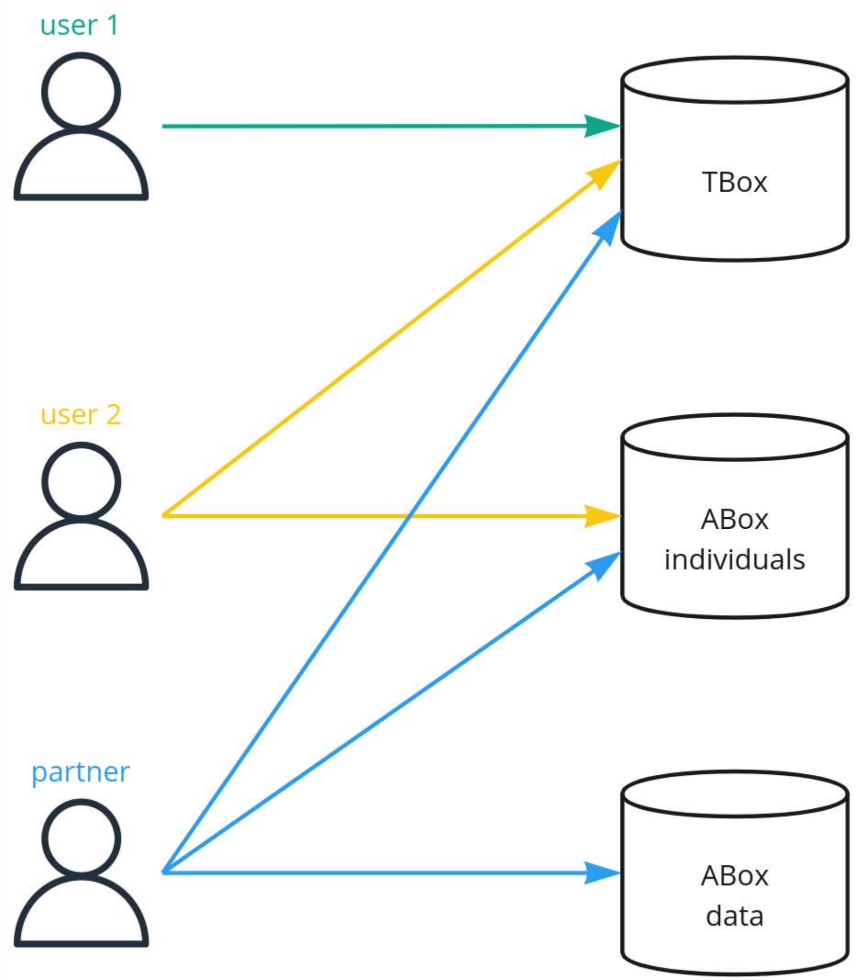
6. Sharing

o Are your data sharable with 3rd parties?

o Do you publish your code/Apps/workflows?

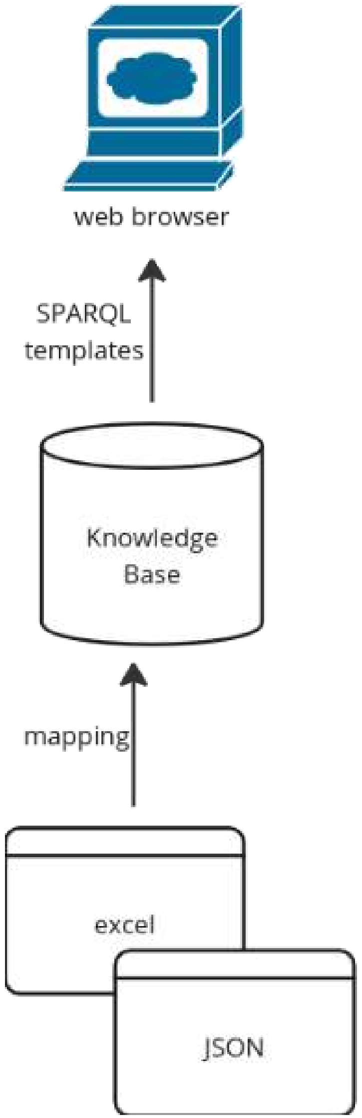
- We identify three distinct types of users, categorized by their access levels:

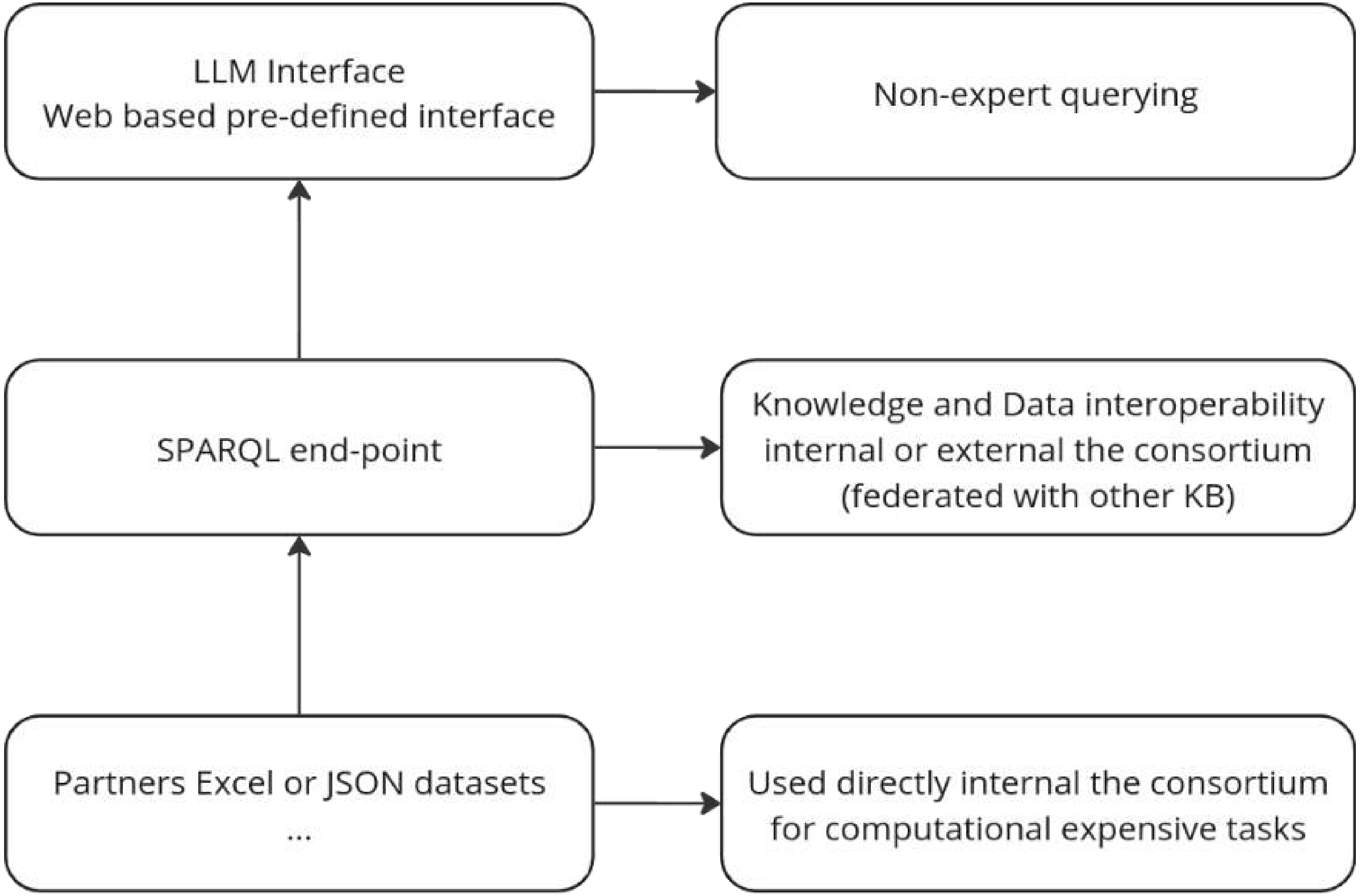
- **TBox-only users:** These users have access exclusively to the TBox, so the connection between entities.
- **TBox and ABox users:** These users can access the TBox and individuals within the ABox.
- **Partners with full access:** These users have access to the TBox, ABox individuals, and the real data.



6. Sharing

- Are your data sharable with 3rd parties?
- Do you publish your code/Apps/workflows?





CoBRAIN Actual Status

We now have:

- Developed 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.

Our Team - 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

Funded by the European Union

Thank you for your attention!

Name: Emanuele Ghedini

Entity: University of Bologna

[redacted]

Funded by
the European Union