

Ontology of Innovation Cases

Approach and Process via Conceptualisation to FAIR Domain Ontologies

Emanuele Ghedini – UNIBO

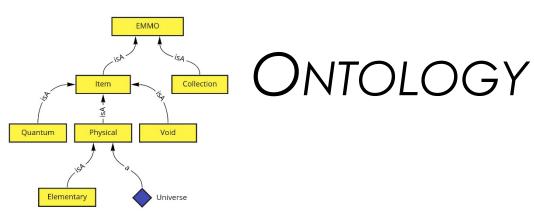


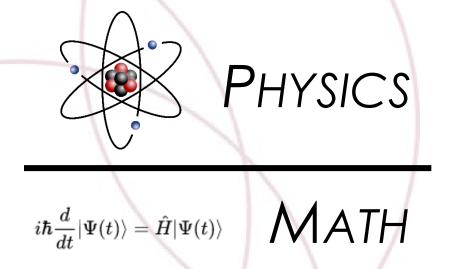
OntoTrans Open Workshop I – 15th March 2022

Ontology As Language

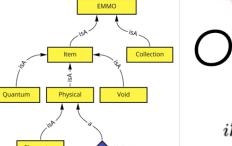


Applied Sciences Knowledge and Data





Ontology As Language



ONTOLOGY

 $i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle$ MATH



Universal Language

(i.e. transcend linguistic differences)

Hard to learn

(i.e. requires adequate training)

Able to Perform Logical Reasoning

(i.e. infer new knowledge from existing one)

Provides Models of the World

(i.e. let us play with abstractions instead of material things)

When It Works, It Can Be Hidden

(i.e. things using it may confine its complexity behind the user interface layer)

Ontology As Language

PHYSICS

MATH

 $i\hbarrac{d}{dt}|\Psi(t)
angle=\hat{H}|\Psi(t)
angle$

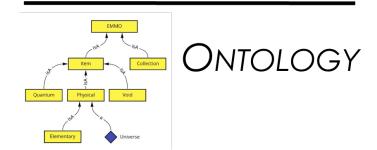
Math has been the key for the highest achievements in Physics

(e.g. Newton, Maxwell, Einstein, Schrodinger, Standard Model)

Ontologies can be the key for a quantum leap in **Applied Sciences** and **Industry** data (e.g. pervasive digitalization, knowledge sharing, Industry 4.0 to 5.0)



Applied Sciences



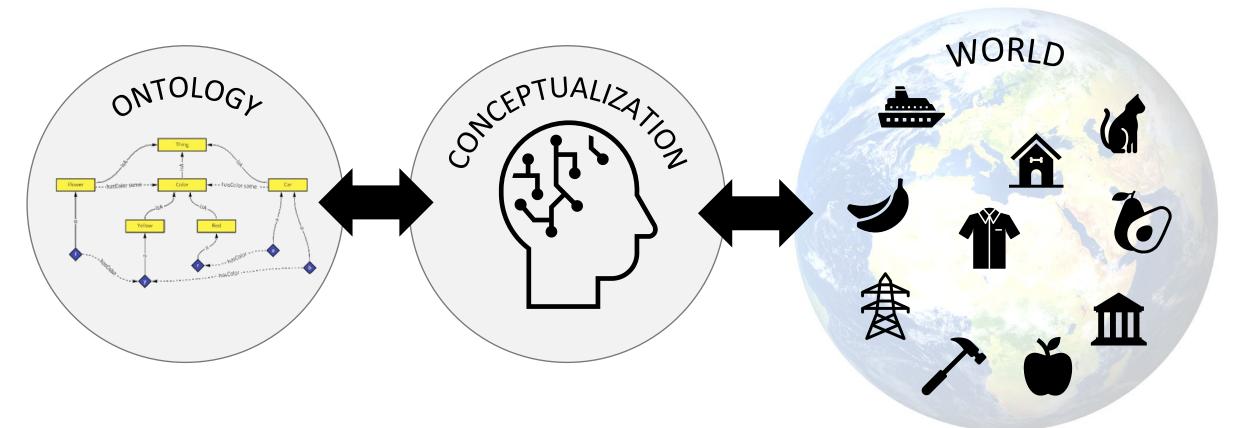
Talking to Machines

Machines understand the logical language of ontologies (e.g. FOL, OWL-2), and already can be used to:

- document data (e.g. dcterms, DCAT)
- infer new knowledge (e.g. resoners)
- support AI (e.g. ontology assisted AI)



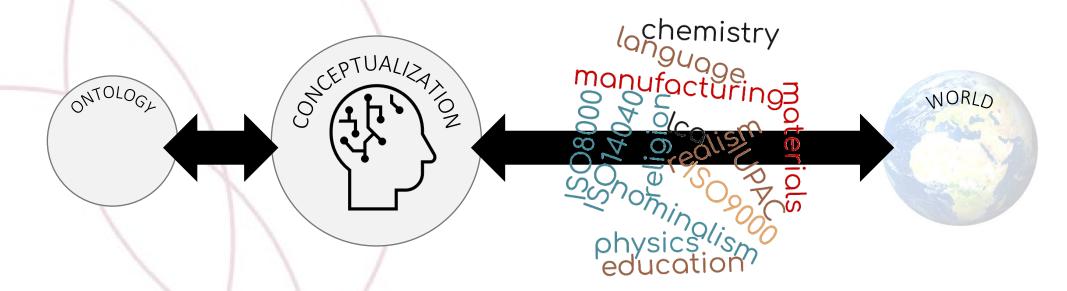
Conceptualisation



Physics uses Quantities (numbers) to model the world Ontologies use Concepts to build a (onto)-logical representation of the world

Conceptualisation

Jumping from Ontology to World requires Interpretation



Several conceptualizations exist for **the same things**, so that almost each human being is going to provide a different definition for a single term. (physics is not affected since it works with quantities, and solve the issue with well defined measurement practices)

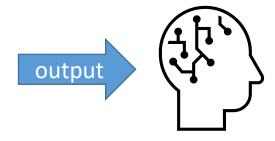
OTRANS

ONT

Conceptualisation

input

Whatever Tool We Create for Our Purposes



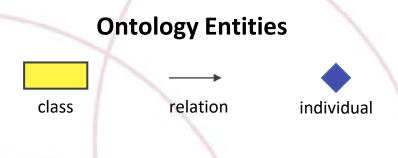
IONTOTRANS

Everything we create to satisfy our needs begins with a **human meaningful input** and a **human meaningful output** (e.g. food, sound, product, action, picture, dataset). e.g. Von Neumann architecture without I/O is a useless machine!

An ontology provides concepts to **semantically categorise such input and outputs**, according to a particular perspective, so that a user can understand what type of entities a particular tool may provide.

e.g. ChatGPT: good for assemble realistic language structures, but no idea about the meaning

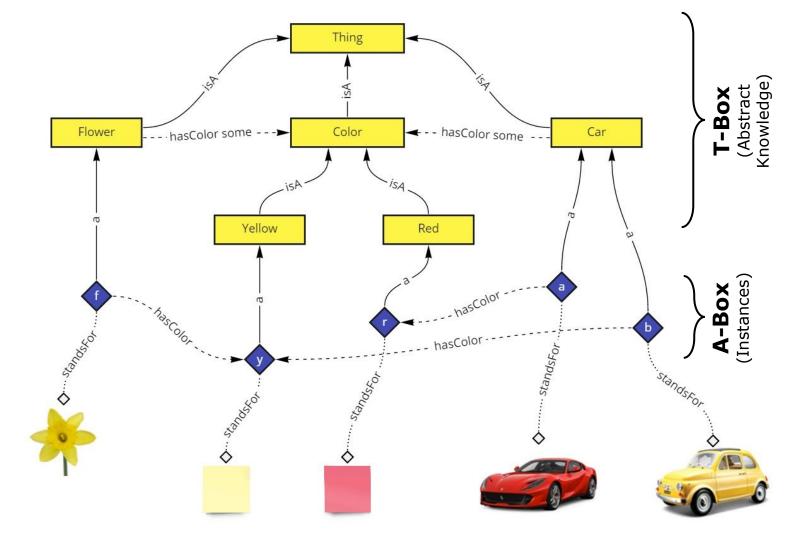
Talking to Machines



Example of axioms: ClassAssertion(:a :Car) ClassAssertion(:r :Color) ObjectPropertyAssertion(:hasColor :a :r)

Real-world objects

No specific ontological commitment about the meaning of 'real' and 'object' in OWL 2. Relying on common sense.



Talking to Machines

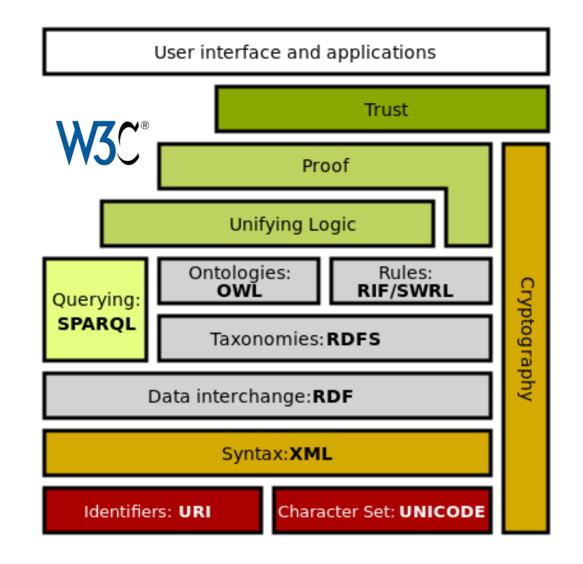
A great boost towards the use of formal ontologies in practice came in the '90 from the **Semantic Web** extension of the **World Wide Web**, thanks to the **W3C** standardization activities.

The objective Semantic Web is **to make Internet data machine-readable**.

Ontologies (in particular formal ontologies) play an important role in the Semantic Web and are placed in the higher levels of the architecture of languages.

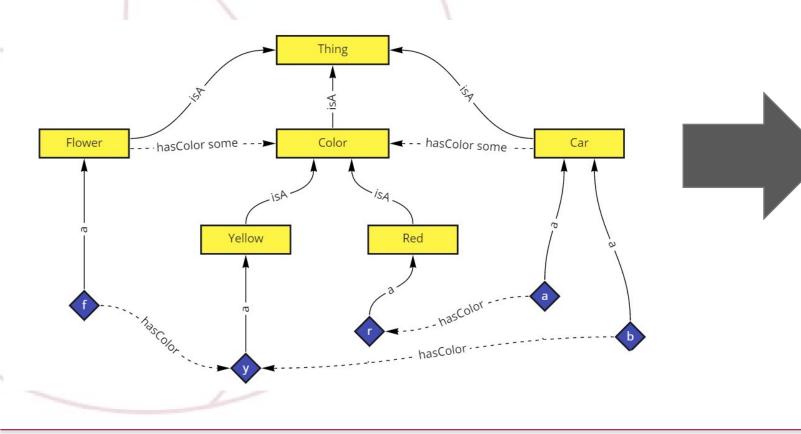
However, the scope and ambition of such ontologies **differ substantially** from the ones coming from the philosophical community.

But an important concept as been introduces: the need for practical applications!



Talking to Machines

An OWL 2 Ontology is formally expressed in a persistent form by axioms declarations following a specific syntax (e.g. ASCII file with Turtle syntax)



@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> . @base <http://www.semanticweb.org/emanuele/ontologies/example> . <http://www.semanticweb.org/emanuele/ontologies/example> rdf:type owl:Ontology . ### http://www.semanticweb.org/emanuele/ontologies/example#hasColor :hasColor rdf:type owl:ObjectProperty ; rdfs:subPropertyOf owl:topObjectProperty ; rdfs:range :Color . ### http://www.semanticweb.org/emanuele/ontologies/example#Car :Car rdf:type owl:Class ; rdfs:subClassOf [rdf:type owl:Restriction ; owl:onProperty :hasColor ; owl:someValuesFrom :Color 1 . ### http://www.semanticweb.org/emanuele/ontologies/example#Color :Color rdf:type owl:Class . ### http://www.semanticweb.org/emanuele/ontologies/example#Flower :Flower rdf:type owl:Class : rdfs:subClassOf [rdf:type owl:Restriction ; owl:onProperty :hasColor ; owl:someValuesFrom :Color 1. ### http://www.semanticweb.org/emanuele/ontologies/example#Red :Red rdf:type owl:Class ; rdfs:subClassOf :Color . ### http://www.semanticweb.org/emanuele/ontologies/example#Yellow :Yellow rdf:type owl:Class ; rdfs:subClassOf :Color . ### http://www.semanticweb.org/emanuele/ontologies/example#a :a rdf:type owl:NamedIndividual , :Car ; :hasColor :r . ### http://www.semanticweb.org/emanuele/ontologies/example#b :b rdf:type owl:NamedIndividual , :Car ; :hasColor :y . ### http://www.semanticweb.org/emanuele/ontologies/example#f :f rdf:type owl:NamedIndividual , :Flower ; :hasColor :y . ### http://www.semanticweb.org/emanuele/ontologies/example#r :r rdf:type owl:NamedIndividual , :Color .

@prefix : <http://www.semanticweb.org/emanuele/ontologies/example#> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

@prefix owl: <http://www.w3.org/2002/07/owl#> .

http://www.semanticweb.org/emanuele/ontologies/example#y
:y rdf:type owl:NamedIndividual ,
:Color .

Facilitate Users' Life **TRANSLATORS ONTOLOGISTS INDUSTRIAL USERS** Human Friendly Graphical Formal Ontology Conceptualisation Template : <http://www.semanticweb.org/emanuele/ontologies/example#> Gprefix oi: <http://www.w3.org/2002/70/owl#s-@prefix oi: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> @prefix xml: <http://www.w3.org/199/02/22-rdf-syntax-ns#> @prefix xml: <http://www.w3.org/X001/X0LSchema#>. prefix rdfs: <htp://www.w3.org/2000/01/rdf-schema#> base <http://www.semanticweb.org/emanuele/ontologies/ex</pre> Car tp://www.semanticweb.org/emanuele/ontologies/example> rdf:type owl:Ontology ### http://www.semanticweb.org/emanuele/ontologies/example#hasColo rdf:type owl:ObjectProperty ;
rdfs:subPropertyOf owl:topObjectProperty rdfs:range :Color . ### http://www.semanticweb.org/emanuele/ontologies/example#Cau :Car rdf:type owl:Class ; rdfs:subClassOf [rdf:type owl:Restriction ; property of real work entities owl:onProperty :hasColor owl:someValuesFrom :Colo ### http://www.semanticweb.org/emanuele/ontologies/example#Colo Property 3 :Color rdf:type owl:Class Property 1 ### http://www.semanticweb.org/emanuele/ontologies/example#Flower Property 2 :Flower rdf:type owl:Class ; rdfs:subClassOf [rdf:type owl:Restriction : output owl:onProperty :hasColor owl:someValuesFrom :Color

Terminology I (how we understand each others...)

<u>Innovation Challenge</u>: the high-level innovation objective addressed by specific cases expressed in natural language (e.g. master the process parameters, agile response to product, market and regulatory requirements)

Innovation Case (or Application): a specific class of cases (e.g. design, manufacturing, sales) for which the challenge applies, involving processes and material objects and expressed in some human readable form (e.g. diagrams, flow charts).

<u>User Case</u>: a specific implementation of an innovation case for the user (e.g. a step in the manufacturing of a specific product, a manufacturing line of a company).

<u>Workflow</u>: data workflow expressed using MODA (or MODA-like) graphs representing the combination of methods for the prediction of properties of our objects/processes of interests (the totality or part of the user case).

ANS

Terminology II (how we understand each others...)

<u>Knowledge Source</u>: everything that can be used to provide new knowledge from existing one. It needs an INPUT (e.g. case set up for simulation, query for a database) and provide an OUTPUT.

<u>Knowledge Generator</u>: a knowledge source that generate new non-existing knowledge (e.g. a model that provides calculated properties for an object, the INPUT is the setup case and the OUTPUT are the properties)

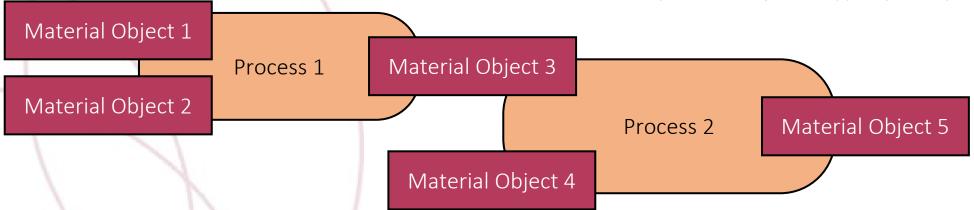
<u>Knowledge Repository</u>: a knowledge source that provides already-existing knowledge (e.g. a database of properties, the INPUT is the name of the substance and OUTPUT is the desired property)

RANS

Innovation Case



EMMO: Causality and not time is the criteria for creating such type of workflows.



An **innovation case** is formalised as an **<u>ontological entities workflow</u>** with the purpose of identifying all the entities that play a role in the case to be referenced by properties and models.

A simplified object/process approach is chosen to facilitate understanding by users without experience in ontologies.

Innovation Case



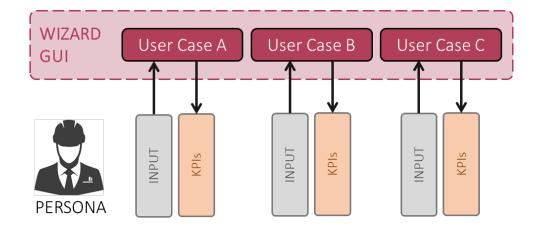
List of all properties we need to deal with, for each innovation case.

Please consider that in the EMMO everything we know about an entity must come as a property, which is a symbolic representation of the object itself defined for a specific class of interpreters within a well defined observation process.

| Property Names | | Data Description (symbolic) | | | | Potential Interpreters | | | Source Data File | |
|----------------|----------------------------------|-----------------------------|------------|-------|--------------------------------|------------------------|-------------|-------------|------------------|--------|
| | Property Description (semantics) | Data Type | Data Range | Units | Reference Physical Quantity | Measured by | Modelled by | Assigned by | Other | Format |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | - |
| | | | | | | | | | | |
| | | | | | | | | - | | |

User Cases

 <u>User Cases are subsets of a specific Innovation</u> <u>Case</u>, expressing the different user approaches to address the Innovation Challenge



- <u>Different User Cases for the same real world innovation case can be proposed</u>, each one focusing on different set of properties or level of details of the objects, according to specific design objectives
- User Cases are represented like the Innovation Case, by <u>selecting objects (or macro</u> <u>objects)</u>, processes and properties of interest
- <u>Properties</u> of real world objects/processes must be defined as problem inputs or KPIs, for each User Case

Workflows

Workflows are intended as data workflow.

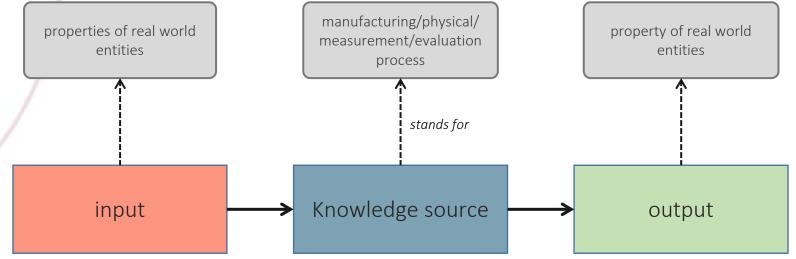
Workflows describe how **properties** are obtained by means of **tools** (e.g. simulation software) or **methodologies** (e.g. measurement).

WIZARD GUI User Case A User Case B User Case C User Case C

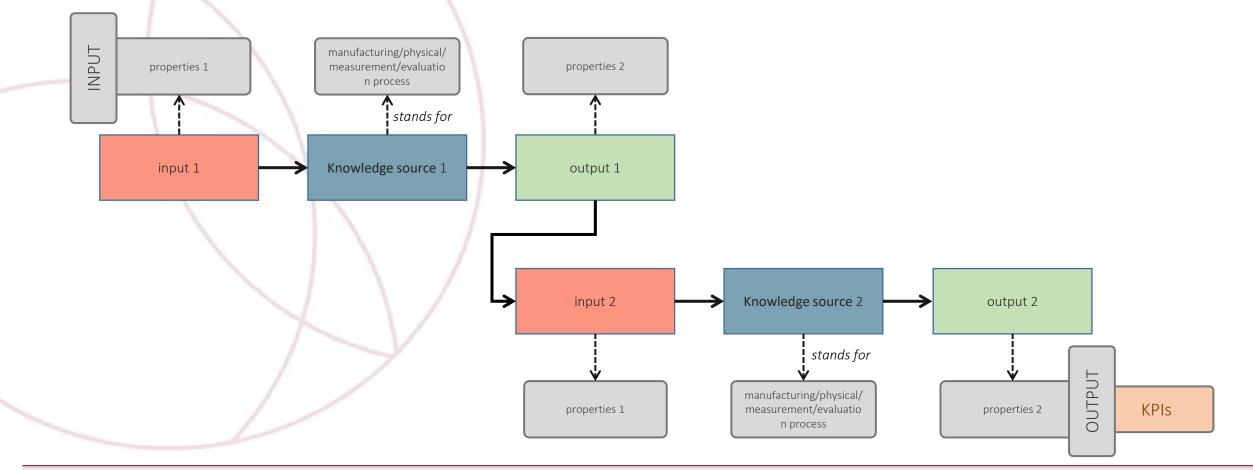
Wokflows are made of a connected **sequence** of **knowledge sources** (e.g. models, databases, experiment).

A knowledge source **stands for** (i.e. substitute, act as, predict the outcome of) a real process.

Workflows must be represented using **MODA-like** graphs, and refers in each step the real world properties and real world processes of the user case



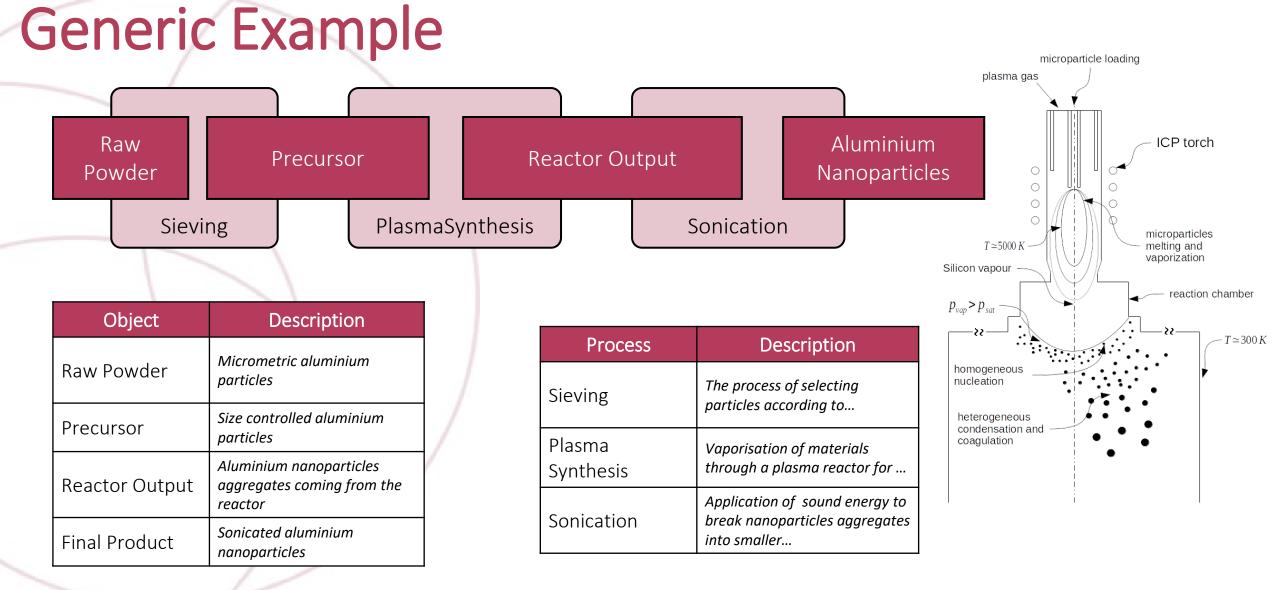
Build one or more potential data workflow connecting data representing properties of real world entities part of the user case



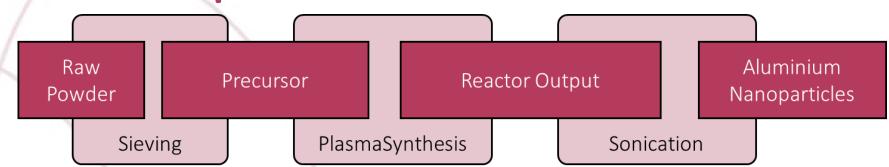
WIZARD GUI User Case A User Case B User Case C User Case C

OntoTrans Open Workshop II – 7^h September 2023

Workflows



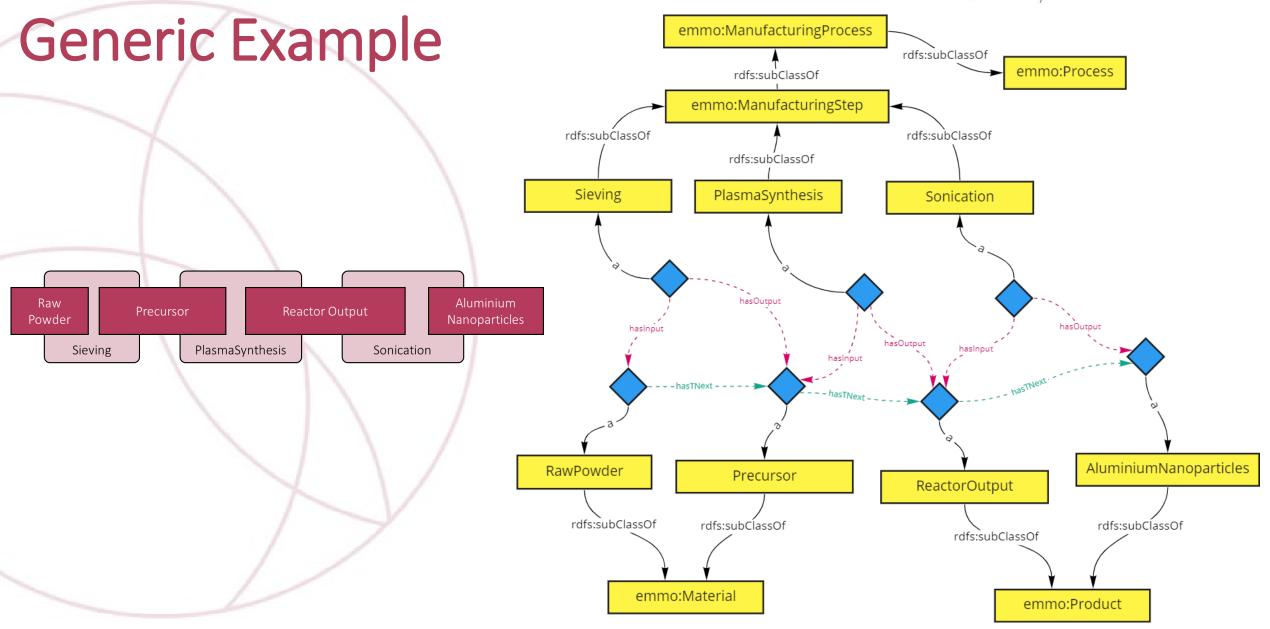
Generic Example



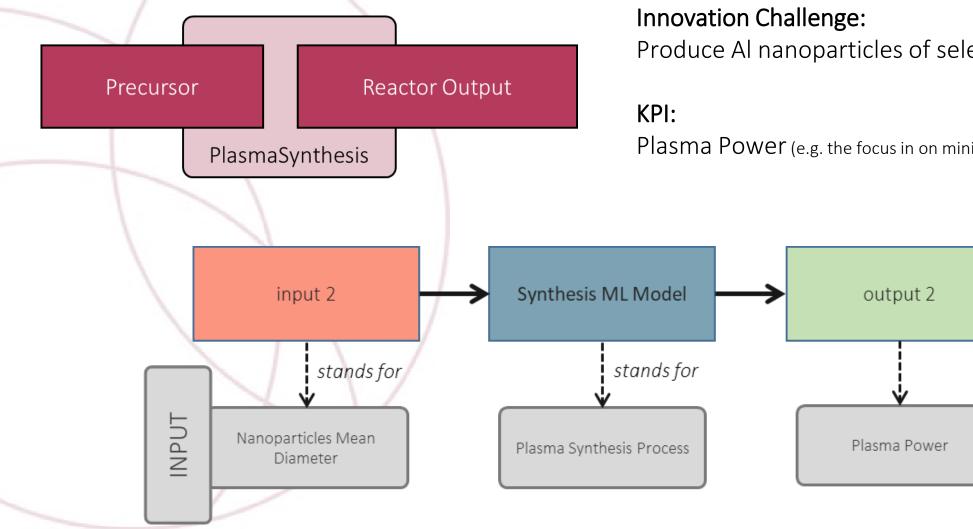
| | Property Names | Property Description (semantics) | Data Description (symbolic) | | | | Potential Interpreters | | | |
|----------------------------|--|--|-----------------------------|------------|-------|--------------------------------|------------------------------|------------------------|-------------|-------|
| Entity | | | Data Type | Data Range | Units | Reference Physical Quantity | Measured by | Modelled by | Assigned by | Other |
| Raw Powder | Powder's mean diameter | The mean diameter of powder's particles. | Real scalar | Positive | m | Diameter | - | - | Seller | - |
| Naw Fowder | Powder's Price | The mass price of the powders | Real scalar | Positive | €/kg | Price per kg | - | - | Seller | - |
| Reactor Output | Nanoparticles mean diameter | The mean diameter of nanoparticles collected after the synthesis. | Real scalar | Positive | m | Diameter | Internal Lab | Synthesis Model | - | - |
| Aluminium Nanoparticles | Aluminum nanoparticles mean diameter | The mean diameter of nanoparticles collected after the sonication. | Real scalar | Positive | m | Diameter | Internal Lab | Sonication Model | - | |
| | Treatment time | The process time. | Real Scalar | Positive | S | Time | Device Timer | - | - | |
| Plasma Synthesis | Plasma Power | The power of the plasma torch in the reactor. | Real Scalar | Positive | W | Power | Reactor instrumentation | - | Operator | |
| Sonication | Sonication power | The power of the sonicator. | Real Scalar | Positive | W | Power | Sonicator instrumentation | Sonication ML model | Operator | |

21

CONTOTRANS



Generic Example



Produce Al nanoparticles of selected diameter

Plasma Power (e.g. the focus in on minimising process power)

KPIS



This was an **easy example of conceptualisation**, based on object, processes, properties, and a simple workflow.

But if we want to represent complex materials, structured datasets, multi-step workflows, we need an **ontology** with an **adequate representational power**.

RANS

Elementary Multiperspective Material Ontology (EMMO)

Started by practitioners in Materials Science in 2018 in order to produce a framework consistent with scientific principles and methodologies.

Developed in a number of projects under the **governance** of the **European Materials Modelling Council** (<u>https://emmc.eu/</u>)

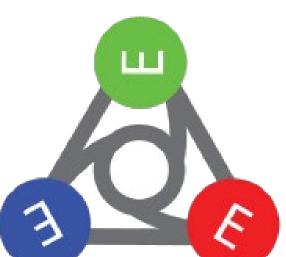
Elementary: it is a fundamental, top-level ontology, while at the same time assuming the existence of objects that cannot be divided further in space and time

Multiperspective: it enables scientists to describe the world by resorting to different views or perspectives, thus providing a high expressive power and versatility

All relationships among objects can be of three types:

- Causal (connectivity)
- Mereological (parthood)
- Semiotic (entities that refer to other entities according to an agent)





ONT

TRANS

IONTOTRANS

Radical Physicalism

We decided to completely embrace a **physics-based description of the world**, starting from the more fundamental and universally accepted knowledge about the **nature**, from Ancient Greek $\varphi \dot{\upsilon} \sigma \iota \varsigma$ (physis): the **Standard Model of Particle**.

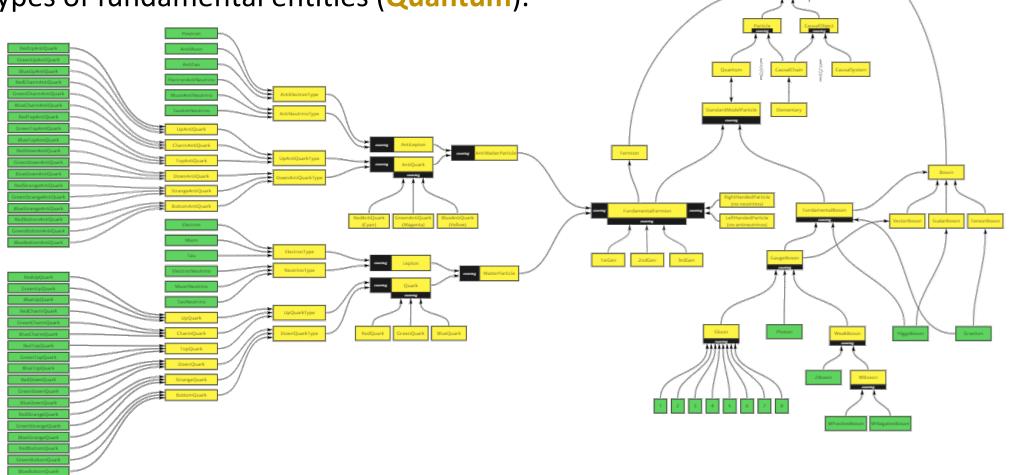
The SMP is the **modern version** of the **ancient philosopher's ontology**: it tells you everything about the world, up to its **fundamental beings**, in Ancient Greek *őντα* (onta)

three generations of matter interactions / force carriers (fermions) (bosons) Ш П ≈2.2 MeV/c² ≃1.28 GeV/c² ≃173.1 GeV/c² ≃124.97 GeV/c² mass 2/3 2∕3 2∕3 0 charge G Η С g U 1∕2 spin 1/2 higgs graviton charm top gluon up $\simeq 4.7 \text{ MeV/c}^2$ $\simeq 96 \text{ MeV/c}^2$ ≃4.18 GeV/c² DUARKS 0 IENSOR BOSONS BOSON -1/3 -1/3 0 S C b 1∕2 1⁄2 1/2 down strange bottom photon SCALAR ≃0.511 MeV/c² ≈105.66 MeV/c² ≈1.7768 GeV/c² ~91.19 GeV/c2 E BOSONS BOSONS $^{-1}$ $^{-1}$ 0 Ζ е Т Ш 1/2 1/2 1/2 Z boson electron muon tau EPTONS <1.0 eV/c² <0.17 MeV/c² <18.2 MeV/c² ≈80.39 GeV/c² 0 ± 1 **GAUG** VECTOR I С О Vμ ντ W Ve 1,6 1/5 1/2 electron muon tau W boson neutrino neutrino neutrino

Standard Model of Elementary Particles and Gravity

Radical Physicalism

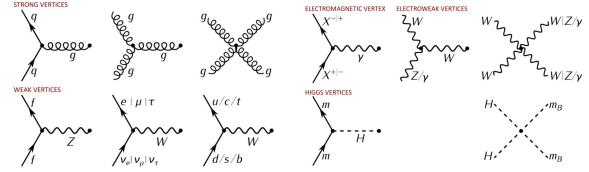
This is the taxonomy of the **EMMO SMP classes**, covering all possible types of fundamental entities (**Quantum**).



LONTOTRANS

Radical Physicalism

Then we had to deal with **physical phenomena** describing how particles interacts. We decided to build a causality theory in order to represent the **Quantum Field Theory** interactions, using **Feynman diagrams** representations.



It was amazing to see the theory grew almost naturally, and how easily the Feynman diagrams rules led to a **causal graph representation** of physical phenomena as **Direct Acyclic Graphs** with a **maximal graph dimension of 4**, totally inline with our **four-dimensional** approach.

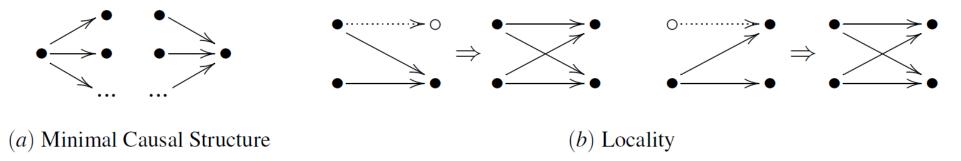


Figure 1. EMMO's Core Naturalistic Commitments (the arrows stand for dC relations)

RANS

IONTOTRANS

EMMO Mereocausality

In order to go beyond the standard model, we decided to go for the fusion of **mereology** and **causality**, creating a new discipline called **mereocausality**, that has been formalised in First Order Logic.

The adoption of **AGEM** theory also led to the definition of **universe**: the **smallest** and the **largest** entities are described using these 13 axioms.

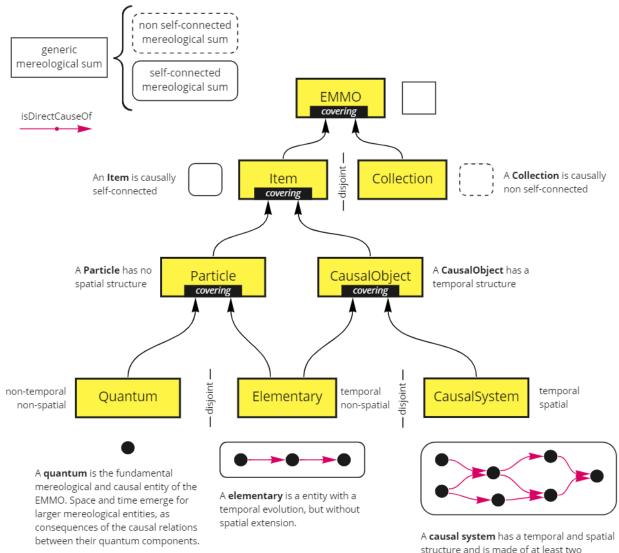
The EMMO mereocausality theory is fully capable of representing everything that exist... but is a bit too detailed for practical usage!!!

```
al P(x,x)
                                                                       (Parthood: Reflexivity)
 a2 P(x,y) \land P(y,x) \rightarrow x = y
                                                                   (Parthood: Antisymmetry)
 a3 P(x, y) \land P(y, z) \rightarrow P(x, z)
                                                                      (Parthood: Transitivity)
 a4 \neg P(y, x) \rightarrow \exists z (P(z, y) \land \neg O(z, x))
                                                                    (Strong Supplementation)
 a5 \exists x(\phi(x)) \rightarrow \exists z(F\langle \phi(x) \rangle(z))
                                                                  (Unrestricted Composition)
 a6 \forall x \exists y (q P(y, x))
                                                                                      (Atomicity)
 a7 \neg C(x, x)
                                                                     (Causality: Irreflexivity)
                                                                     (Causality: Transitivity)
 a8 C(x, y) \land C(y, z) \rightarrow C(x, z)
 a9 C(x,y) \rightarrow dC(x,y) \lor \exists zw(C(x,z) \land dC(z,y) \land dC(x,w) \land C(w,y))
                                                                                   (Discreteness)
a10 C(x,y) \rightarrow Q(x) \land Q(y)
                                                                          (Quantum Causality)
all dC(x,y) \to \exists z((dC(x,z) \lor dC(z,y)) \land y \neq z \land x \neq z)
                                                                 (Minimal Causal Structure)
a12 dC(x, y) \land dC(x, z) \land dC(w, y) \rightarrow dC(w, z)
                                                                                         (Locality)
a13 ITEM(u)
                                                                          (Connected Universe)
```

IONTOTRANS

EMMO OWL Top Level

EMMO represent world entities as a causal **network** (a Direct Acyclic Graph) of quantum entities whose types and interactions are governed by the Standard Model of Particles.



Spatial and Temporal Relations

The EMMO defines **spatial and temporal relations** according to the their mereocausality structure.

The **rules** derived from Feynman diagrams are paramount for the creation of such relations i.e., spacetime emerges thanks to them.

Graphical Mereological Relations

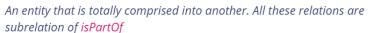
Express the mereological relations between concepts through a 4D graphical representation using the following schemes representing how objects and processes can relate each others. Each case is labelled according to how the grey boxes relate to the white ones.

Proper Overlap

Two entities that share some of their parts, without being one part of the other. All these relations are subrelation of *isProperOverlapOf*.

| isAddedTo | isOutputOf | affects | contributesTo |
|---------------|------------|------------|----------------|
| isRemovedFrom | isInputOf | partakesTo | participatesTo |

Part Of



| - space | isConstituentOf | isConstitutiveProcessOf | isProperParticipantOf | isSubprocessOf |
|---------|-----------------|-------------------------|-----------------------|----------------|
| time – | | isBehaviourOf | isStatusOf | isStageOf |

Mereological relations can also be expressed with **red** arrows

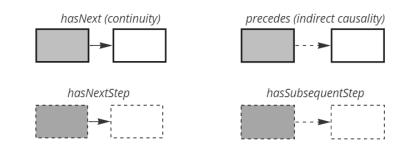
Graphical Causal Relations

Express the causal relations between concepts though a 4D graphical representation using the following relations, expressed by black arrows.

IONTOTRANS

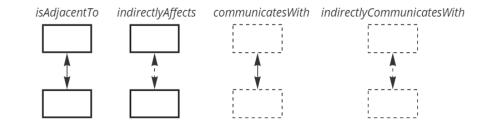
Temporal Causality

Describes the causal relations between entities that unfold in time.



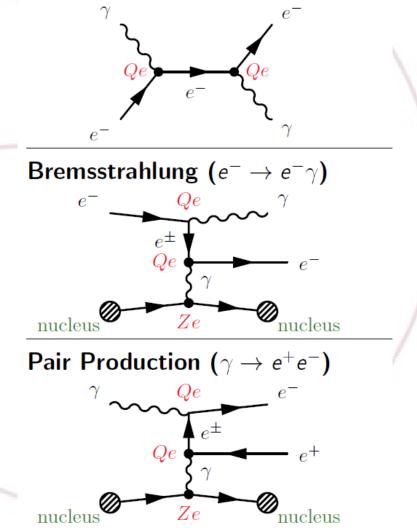
Spatial Causality

Describes the causal relations between entities that unfold in space (symmetric).

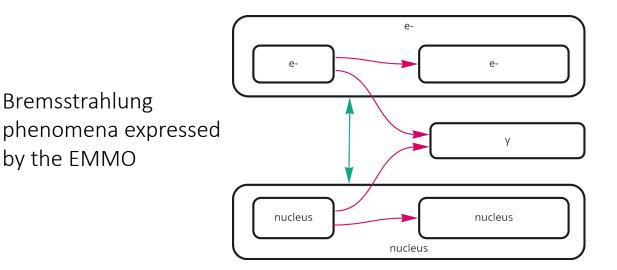


Feynman Diagrams

Compton Scattering ($\gamma e^- \rightarrow \gamma e^-$)



Interactions between quantums are described by **Standard Model Theory** (QFT), expressed usually through Feynman diagrams, enabling the highest possible ontological detail and objectivity in the representation of materials.



Classical Fields

Fusion Reactor

Interactions are represented as mediated by **real particles** (i.e., a causal chain) so that it is possible to apply **non-quantum** approaches (i.e., classical or relativistic). A **classical field** is represented as a **stream of real particles**.

Plasma



IONTOTRANS

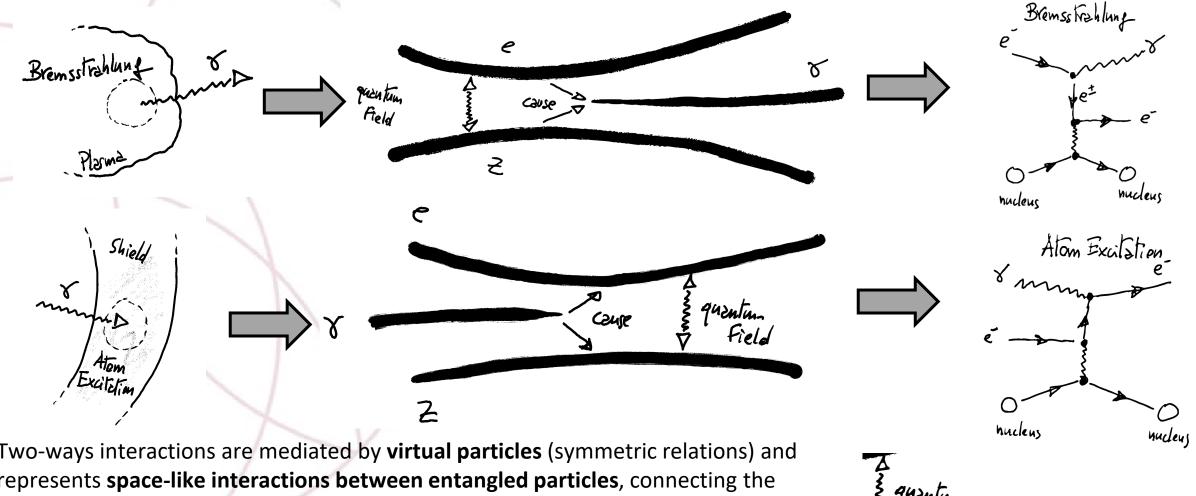
Bremsstrahlung

Plarma

Shield

Two-ways interactions are mediated by virtual particles (symmetric relations) and represents **space-like interactions between entangled particles**, connecting the entities used to build the Hamiltonian for a quantum mechanical representation. A quantum field is represented as a symmetric relation.

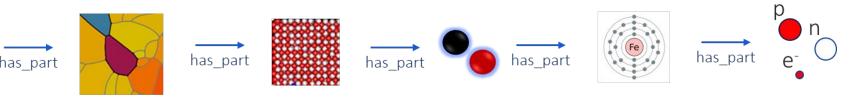
Quantum Fields



Representing the World



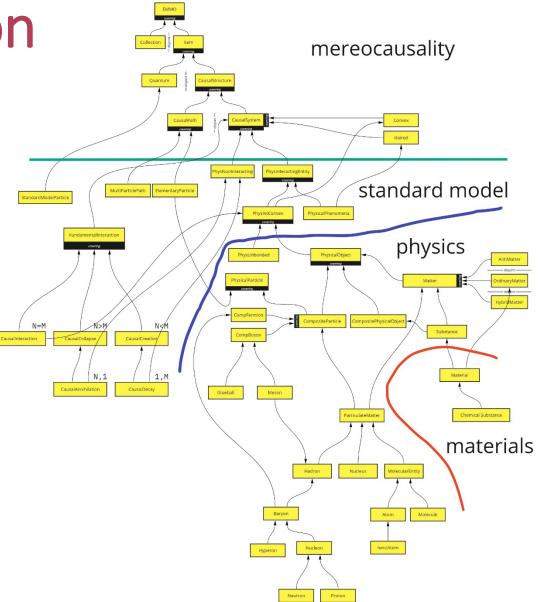
The EMMO needs to represent all the entities between SMPs up to the universe. How do we deal with e.g., composite particles, atoms molecules, materials, phases, solids, mixtures?



Multiscale Representation

The EMMO uses **mereocausality** to classify entities according to their **structures**, distinguishing between concepts like **physical phenomena** (e.g., scattering) or **objects** (e.g., atom).

This provides a **seamless multigranular representation of world entities**: exactly what you would like a material ontology to provide!



From Elementaries to Materials

Why we need a so deep fundamental **level of detail** in the description of materials?

Because we want:

- to speak the language of material scientists (e.g. physics, chemistry)
- to enable unambiguous representation of world entities ready for been translated into one of the existing physicsmathematical models (e.g. quantum mechanics)
- to enable multi-scale representation of world entities
- to provide a strong foundation for material-based taxonomies

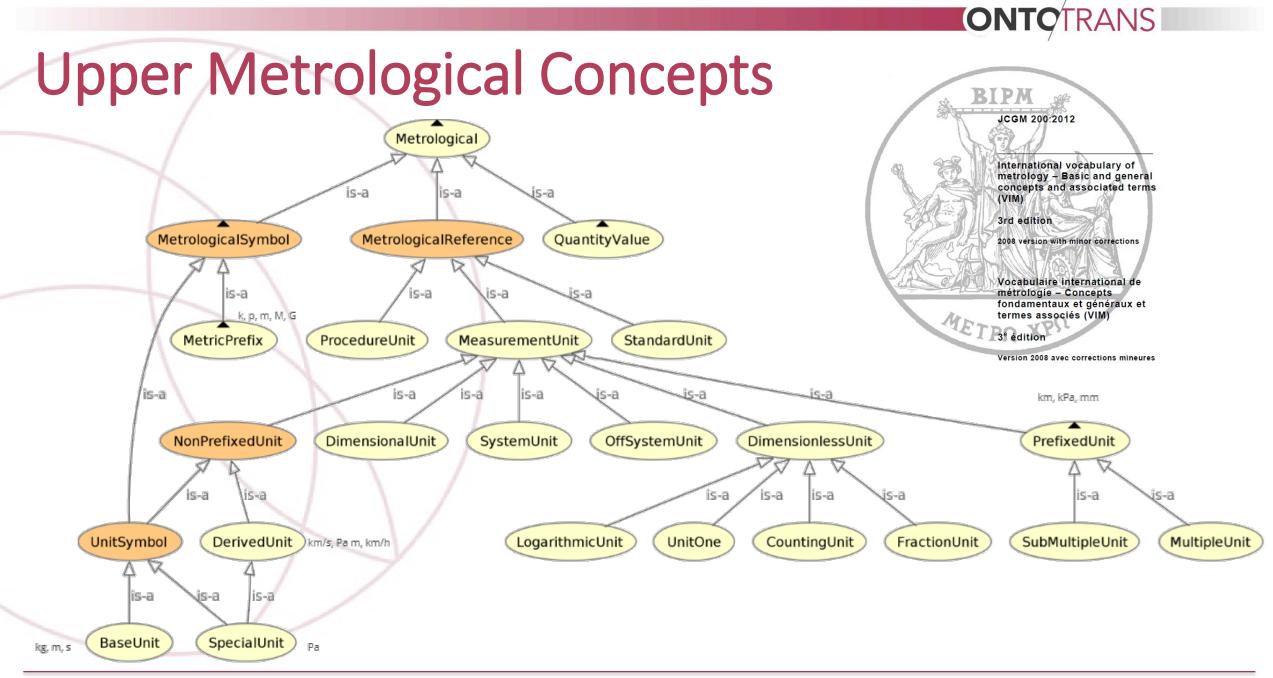
RANS

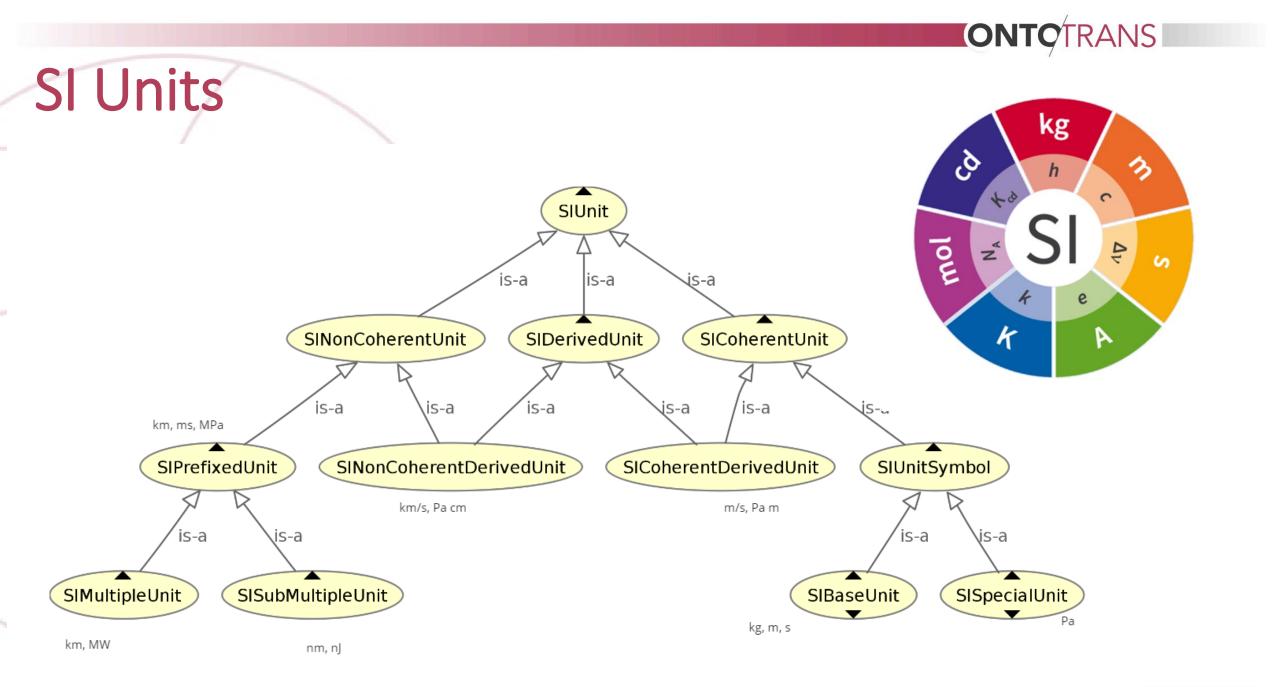
Physical Quantities

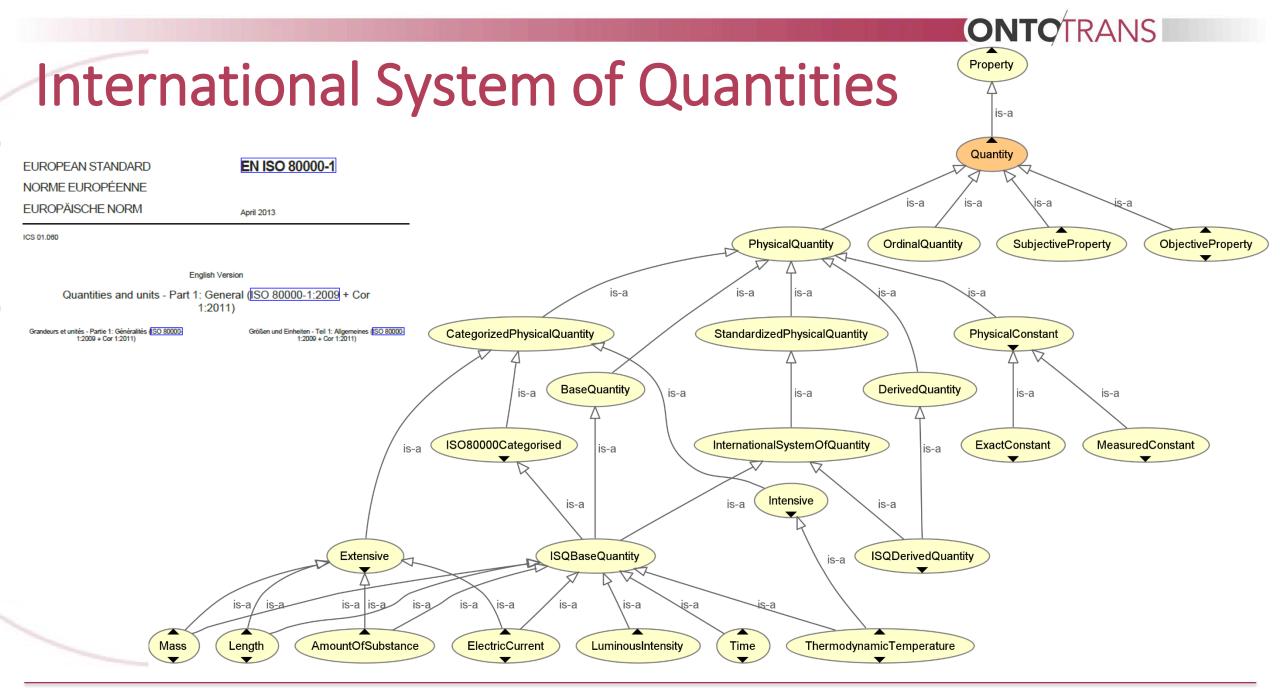
We wanted to build a representational system able to embed ISO 80000 International System of Quantities, the Vocabulary of Metrology, and the SI International System of Unit in a way that is clearer and more expressive than existing ontologies (e.g., QUDT).

To do so we **separated the semantic from the syntactic part** in the representation of **physical quantities**.

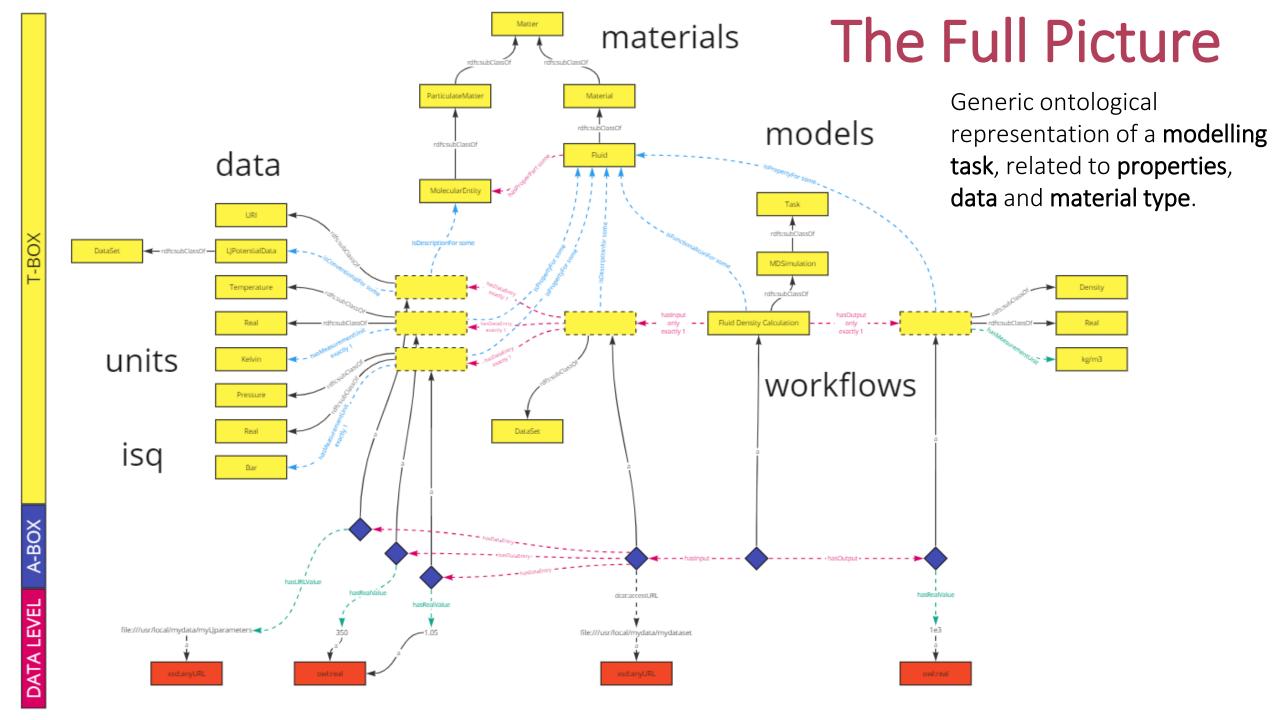
RANS







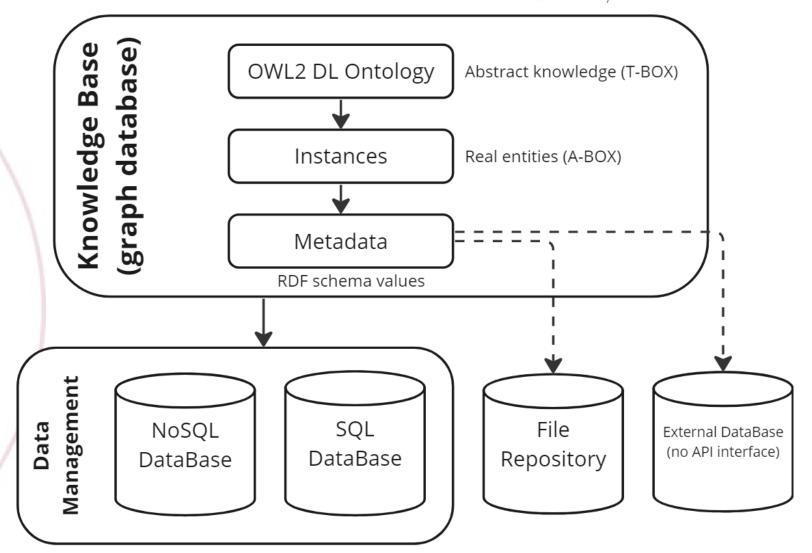
CONTOTRANS **Physical Quantities SEMANTIC MEANING** (e.g., length) **UNIT SYMBOL** PREĖIX **DATA VALUE MEASUREMENT UNIT**



Knowledge Management

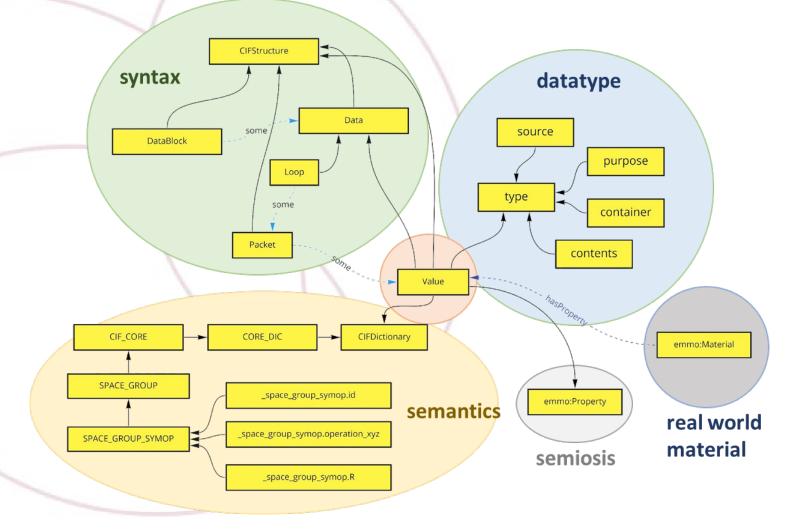
We can create a **Knowledge Base** based on EMMO, that is built upon existing **Data Management** solutions, implementing **data federation** in practice.

Ontologies are a great way to make your data (ore better, your knowledge) **FAIR**.



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.

EMMO and **CIF**



The acronym **CIF** is used both for the **Crystallographic Information File** and for the **Crystallographic Information Framework**, a broader system of exchange protocols based on data dictionaries and relational rules expressible in different machinereadable manifestations.

The EMMO CIF ontology maps all concepts of CIF-DDL and CIF-CORE and provides a multiperspective approach to represent CIF data and to connect them to real world materials. Conclusions

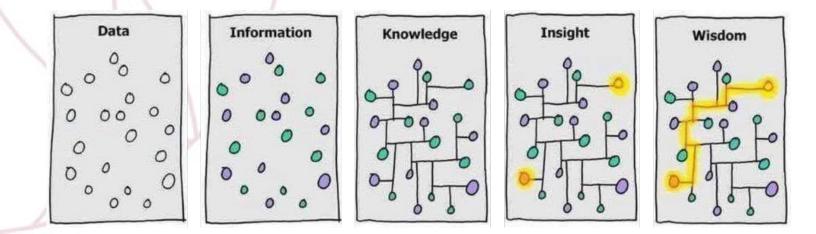
- The use of ontologies as **human-to-machine** and **machine-to-human** interfaces requires an approach that facilitates the participation of industrial users in the **conceptualisation step** (i.e. OntoTrans template)
- Human Translators are needed to guide the industrial users and formalise the user case in ontological form.
- The EMMO is used to build ontology modules to represent all the steps of Innovation Case to **design**, **execution** and **documentation**.
- Ontological representation is used to populate a knowledge base (OntoKB) that can be navigated by agents to harvest existing data, or used as training set for AI approaches towards automated decision making or workflow design.

RANS

In Layman Terms...

The ontology is expected to <u>represent</u>, to <u>document</u> and to <u>connect</u> together:

- the innovation and user case (i.e. the things: object and processes)
- their properties (i.e. the data that describes the things)
- the tools used to generate (e.g. models, characterisation) or retrieve (e.g. databases) the required properties (i.e. the knowledge sources)
- the data workflow used to generate the desired properties (i.e. the KPIs)



TRANS

ON1







The OntoTrans project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 862136.