Top Level Ontologies: desirable characteristics in the context of Materials Science

SeMatS

2024

Pierluigi Del Nostro[1], Jesper Friis[2], Emanuele Ghedini[3], Gerhard Goldbeck[1],Daniele Toti[1] and Francesco Antonio Zaccarini[3]

1 | Goldbeck Consulting Limited, Cambridge, United Kingdom

2 | SINTEF AS, Trondheim, Norway

3 | Alma Mater Studiorum - University of Bologna, Bologna, Italy

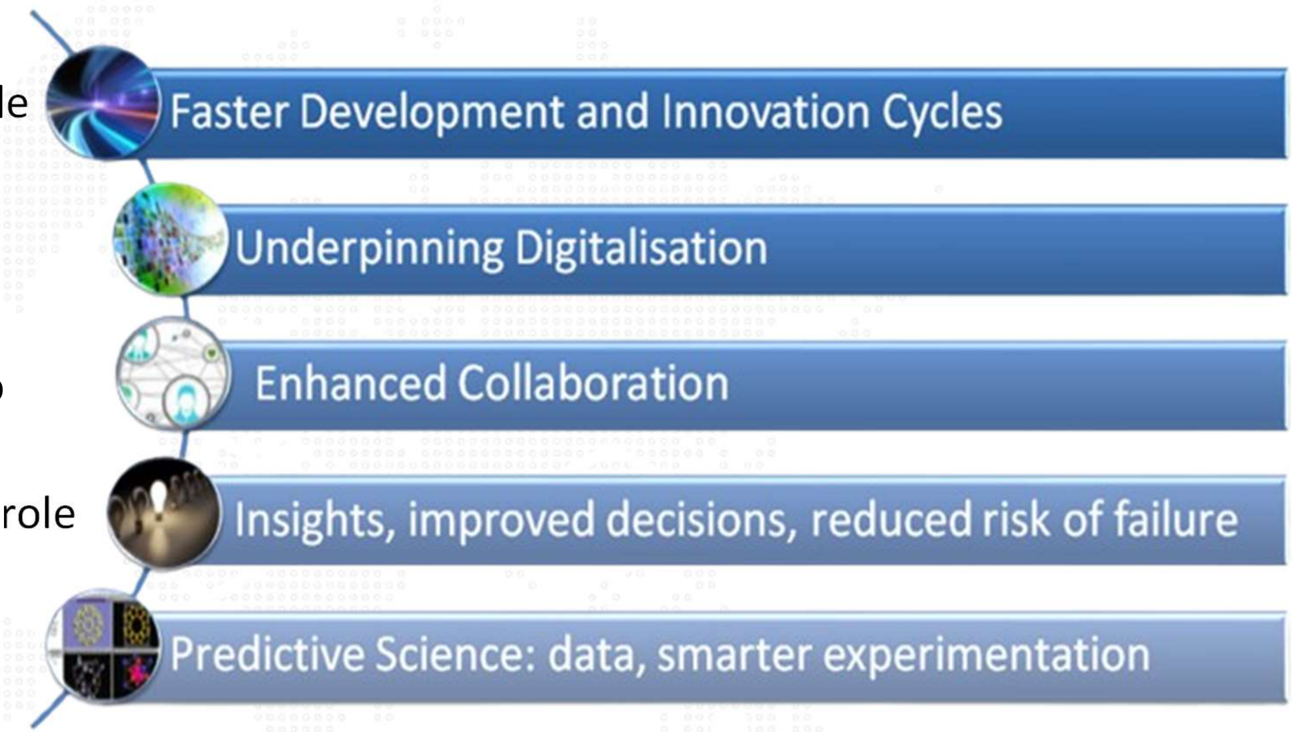
GOLDBECK CONSULTING

SINTEF

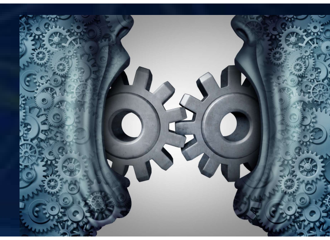
Amsterdam | 17/09/2024

Top Level Ontologies for Materials Science

- **Semantic technologies** play a core role for driving **innovation**.
- Materials development draws on a **plurality of subdomains and data resources**, making the adoption of a **Top-Level Ontology** recommended to ground **interoperability**.
- Top-Level Ontologies also play a core role when it comes to **content-based representations**, as they stand for a particular **worldview**.

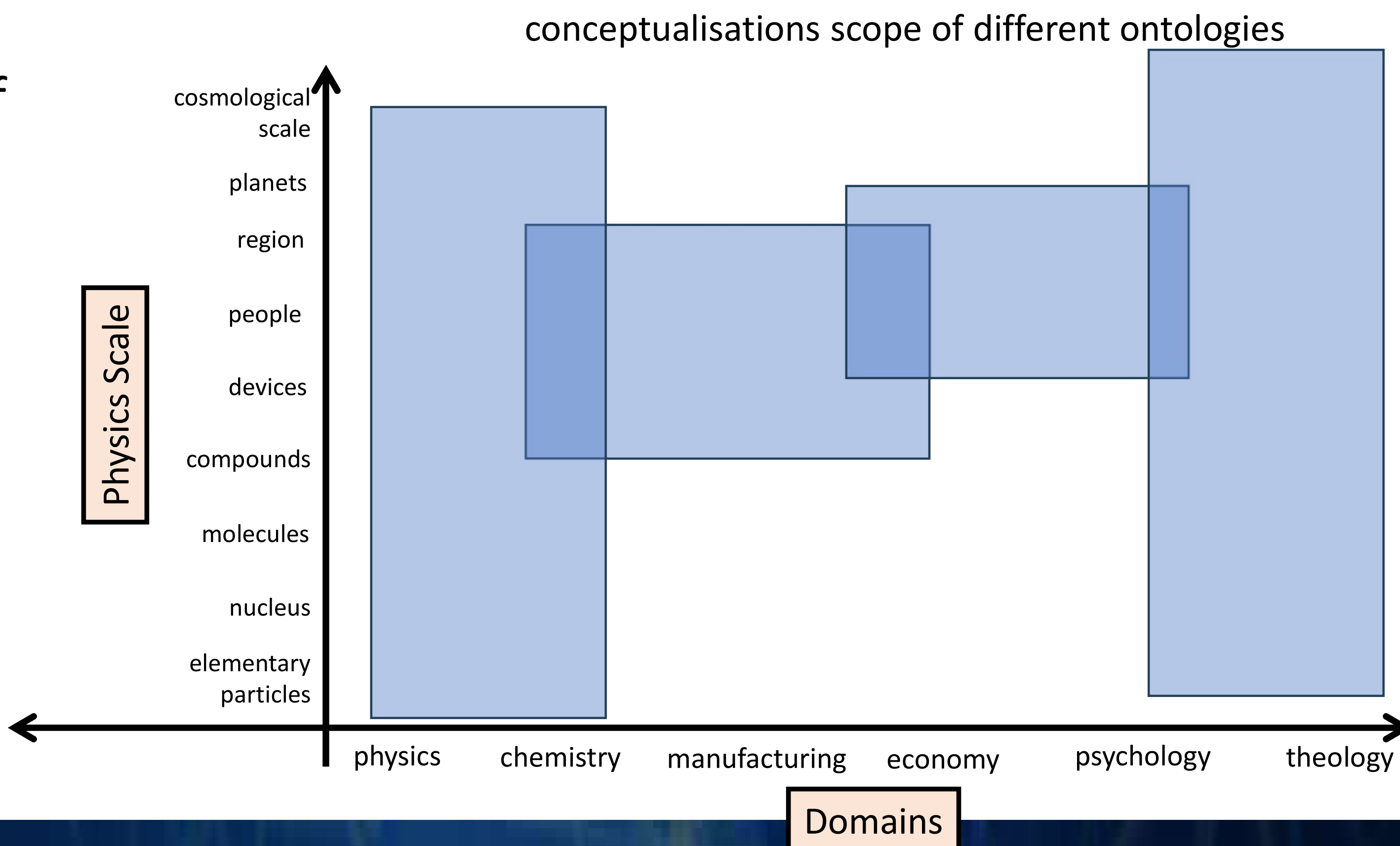


TLOs and their conceptualisation scope



Ontologies (incl TLOs) display varying degrees of suitability in different contexts and given different pragmatic goals/use cases.

Investigate the **pros and cons of different choices** regarding the **needs of Materials Science and cognate domains (MSE)**



Methodology: following ISO/IEC 21838-1

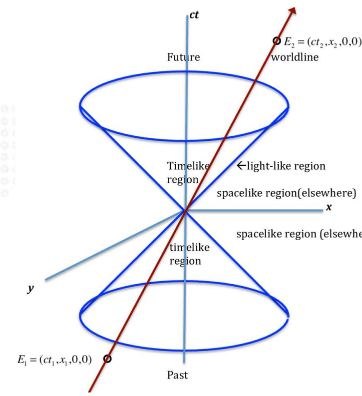
Information technology — Top-level ontologies (TLO) — Part 1: **Requirements**

- A list of general requirements for TLOs
  - **Conceptual/representational requirements:** topics to be covered by a TLO.
  - **Formal requirements:** formal languages etc
- Leave room for **alternative design choices/solutions**
- Evaluate the pros and cons of **choices relative to domain requirements.**
- Consider the **scientific world-view and** and prototypical MSE use-cases.

ISO 21838-1:
requires explanations of how topics are covered by the TLO.

Very different approaches to specific topics are possible

- *Location:*
  - Objects and processes/events can be co-located in **DOLCE**
  - Different entities cannot be co-localized: **BFO**
- *Time and Change*:
  - 4-dimensionalist: **ISO 15926**
  - Not strictly 4-D: **DOLCE** and **BFO**.
- *Not every topic has to be covered directly by respective entities*
  - TLO may reject the relevant entities in their domain, recovering the related expressiveness through other means:
  - **E.g.:** *Actuality and Possibility*: possibilia are not allowed in **BFO**.



Selected Conceptual Desiderata: Qualities

4.4.6.9 Qualities and other attributes

How does the ontology deal with qualities and other attributes? NOTE 'Attribute' here is meant to include what are sometimes referred to as properties, features or characteristics.

Properties as **intrinsic** characteristics of entities: has practical and theoretical limits.

- Science distinguishes between theory and observation.
- We only know what we measure, which is limited by uncertainty and theoretical assumptions.

Properties as signs attribution (nominalist approach), with potential choices:

- Relativization of properties to observations/models à la Masolo, Benevides & Porello (2018).
- Closer to scientific method: observation combined with theory/models to declare a property

Selected Conceptual Desiderata: Quantities

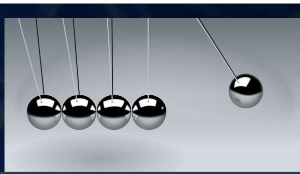
4.4.6.10 Quantities and mathematical entities

How does the ontology deal with quantitative data and with mathematical data and theories?
How does the ontology deal with units of measure?

- Quantities (and related qualities!) should be explicitly grounded in specific measuring or calculating procedures, following ISO 10303-45:201 and metrology.
  - A “property” is defined by relating the result of a particular measurement process to the measured object.

- The way in which a quantity or quality was determined should be supported in an ingrained way
  - Avoid assumptions about the meaning of data:
    specify “what is measured and how that relates to property”

Selected Conceptual Desiderata: Causality



- Laws and causal relations play a core role in scientific reasoning.
- In industrial operations, workflows can be understood in causal terms.

4.4.6.13 Causality
— How does the ontology deal with causality?

- Not well covered by ISO-standard - TLOs, mostly relying on spatiotemporal relations.
- Conflicting interpretations are common:
  - E.g. BFO-based ontologies define the nature of the cause as either a material entity or a process.*
- MSE is best served by an explicit causal relation playing the role of laws.
  - A relation is preferable to axiomatic constraints: a more organized structuring of specifications etc
- “Causation” covers a family of notions, including counterfactual and productive
  - While both counterfactual and productive notions are employed in scientific reasoning, the former appear to be more appropriate in disciplines more closely related to social sciences.
  - A productive notion with a physical interpretation is thus recommended, possibly capable of supporting extensional identity criteria.
  - Conserved quantities approaches to causation offer various benefits.

* Sawesi, S., Rashrash, M., Dammann, O., 2022. The Representation of Causality and Causation with Ontologies: A Systematic Literature Review. Online J Public Health Inform 14, e4. https://doi.org/10.5210/ojphi.v14i1.12577

Proposed Desirable Characteristics: Summary (1)

Conceptual Requirement	Desirable Design Choice
Space & Time	- Unitary treatment - Separation of measured time and ontological constraints
Qualities and Quantities	- Providing means to represent the subjectivity of observation - Identification of a property with specific measuring or calculating procedures, in accordance with ISO 10303-45:201 & metrology
Classifications and Mathematical, Informational, Social, & Mental Entities	No ontological commitments or strictly reduced to/grounded in Physical Entities
Parthood, Location, Constitution, Time and Change, and Events and Processes	Favouring "reductionist" approaches and extensional criteria of identity: - Adoption of an Extensional Mereology - Identification of entities and location (supersubstantivalism/relational theory of space-time) - Perdurantist approach to Perdurance in time - No substantial ontological distinction between events/processes and objects
Causality	- Direct support through a relation or axiomatic constraints - Adoption of a productive notion with a physical interpretation
Scale and granularity	- Direct support for the representation of the same entity at different scales - Reductionistic approach taking into account scientific pluralism

Further Desirable Characteristics: Summary (2)

- **Domain Neutrality**
  - Design choices are often mutually exclusive, and do not allow for an informative, maximally general alternative.
    - Broad applicability/interoperability not compromised: it is usually unnecessary to represent peripheral domains' knowledge to high degrees of specificity, and alignments are an option.
  - Conceptualizations ought to be shared. The endorsed concepts should be understandable by MSE practitioners, i.e., they have to be rooted in state-of-the-art sciences and related to golden standards recognized by the MSE community.
    - Hard choices due to scientific pluralism – necessity of a solid framework.
- **Formal Requirements**
  - MSE-oriented TLOs should have lightweight machine-readable implementations compatible with less expressive OWL profiles (EL/RL/QL) supported by triplestores
  - An efficient modular architecture, allowing users to employ only portions of the TLO, is paramount.
- Desiderata relative to versioning and maintenance to be explored in the future (can affect architectural choices).

Concluding Remarks

- The EMMO (Elementary Multiperspective Materials Ontology) was developed to align with the individuated ontology design characteristics.
- The individuated additional desirable characteristics still allow for different solutions.
  - The core principles involve the adoption of extensional, scientifically or formally-grounded criteria of identity across the board, with the introduction of ingrained support for the representation of properties in line with metrological principles.
  - Causation and scale & granularity have been individuated as especially important topics a TLO tailored for MSE should focus on.
- Our aim has been to make **design choices more explicit**, and on that basis to improve **knowledge representation in MSE** and related industrial sectors.

Acknowledgements

The European
Materials Modelling Council

This work has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreements Numbers: 953167 (OpenModel), 862136 (OntoTrans), 952869 (NanoMECommons).

Ontology Ecosystem

Open Translation Environment

Harmonisation of characterization protocols

Material Modelling Innovation Platform

Open Simulation Platform

Materials Modelling Marketplaces

Industrial Data Marketplace