Introduction to the VIMMP ontologies

Martin Thomas Horsch, Silvia Chiacchiera, Michael A. Seaton, and Ilian T. Todorov,
UK Research and Innovation, STFC Daresbury Laboratory, Daresbury, Cheshire, UK

Daniele Toti and Gerhard Goldbeck
Goldbeck Consulting Ltd, Cambridge, UK

(date: 2nd January 2021)

Disclaimer: This document is intended for joint dissemination with the ontologies developed by the Virtual Materials Marketplace (VIMMP) project. The content of its first version was included in **VIMMP project deliverable 1.4** and made available to the public through CORDIS; the material (ontologies and this document) can be reused and further developed under the terms and conditions of **LGPL version 3**, provided that proper attribution is made to the original authors, any additional contributors, and the VIMMP project consortium. VIMMP is funded from the EU's Horizon 2020 research and innovation programme under grant agreement no. 760907. **The authors retain the right to publication and dissemination by additional means** [Hor21], following procedures and regulations in line with the terms and conditions from the VIMMP consortial agreement.

1. Summary and overall structure

Digitalization of industrial process and product design and industry 4.0 enhanced manufacturing are tied to innovations in process data technology [Dep18, Mai17, Mal14a]. In this context, semantic technology facilitates the integration of data with a diverse and heterogeneous provenance into coherent frameworks [Gyg20]. By combining multiple source data sets, repositories, or research data infrastructures, simulation results can be evaluated and assessed for consistency [Ste19]. Semantic interoperability is a precondition for the reliable interaction between digital infrastructures through which services are provided and data are exchanged. Metadata standardization at the level of semantics ensures that information on objects can be ingested, extracted, and communicated in a mutually agreed way, facilitating the implementation of FAIR data stewardship within a distributed, heterogeneous semantic-web framework. Ontologies expressed in the Web Ontology Language (OWL) and OWL Description Logic (OWL DL), using formats such as the Terse Triple Language (TTL), are a mechanism for formalizing the required semantic standards in a machine-processable way [All11, Baa17]. Interoperable frameworks (ecosystems), with a large number of separately developed services, software components, and platforms, need to integrate the semantic space pertaining to multiple domains of knowledge – characterized by domain ontologies – into a single coherent formalism.

To enable semantic interoperability with external services and platforms, the VIMMP project has developed a system of marketplace-level domain ontologies, *cf.* Fig. 1.1, supporting the ingest and retrieval of data and metadata at the VIMMP marketplace frontend; these ontologies are expressed in OWL2 using TTL notation [Hor20a]. It is the purpose of the present deliverable to summarize this work and its outcome. Internally, VIMMP uses the marketplace-level domain ontologies as a part of its approach to data management, underlying the interactions with users at its frontend. To coordinate these developments with the community and the ecosystem of platforms developed from related projects funded from the NMBP area of the Horizon 2020 research and innovation programme, VIMMP contributes to the activities of the European Materials Modelling Council (EMMC), particularly the EMMC focus area of digitalization, and it employs the European Materials and Modelling Ontology (EMMO) as a top-level ontology [Fra20].

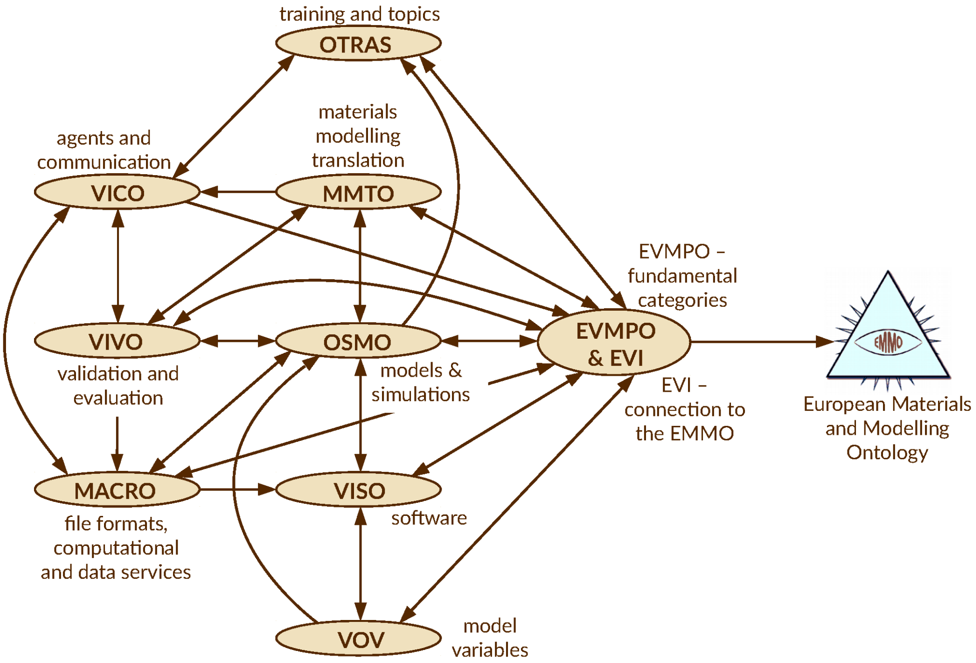


Figure 1.1. Ellipses: Ontologies developed by VIMMP. Triangle: EMMO, the employed top-level ontology. An arrow signifies that an ontology refers to concepts or relations from another ontology.

Thereby, an interoperability framework is established that extends previous EMMC standardization efforts, including the Review of Materials Modelling [Deb17], the EMMC Translation Case Template [Emm17], the EMMC Translators' Guide [Hri19], and the MODA metadata standard for simulation workflows [Cen18]. Providers will have the possibility to choose the depth at which any provided services and tools implement the proposed common semantics, as the deeper the implementation the better the interoperability will be with other services [Hor20a]. In particular, the European Virtual Marketplace Framework (EVMF), established by the joint work of the MarketPlace and Virtual Materials Marketplace (VIMMP) consortia in coordination with the EMMC, is open to participation by any interested provider, translator, or end user of services in materials modelling. The EVMF is entirely based on transparent and openly accessible specifications, relying on the EMMO at the top level;[1] the present ontologies are accordingly released as free software under the GNU Lesser General Public License (LGPL) version 3. By creating an open ecosystem on the basis of community-governed interoperability standards, a variety of projects, many of which (including VIMMP, MarketPlace, and OntoCommons) are funded from Horizon 2020, contribute to a system of platforms and infrastructures that will support the uptake of materials modelling solutions by industrial research and development practice.

[1] Any references to the **European Materials and Modelling Ontology** [Fra20] in the present document, and in the described VIMMP ontology versions at delivery time (June 2020), correspond to **EMMO version 1.0.0 beta dated 14th November 2020**; *i.e.*, the most recent version of the EMMO that was available in time to be taken into account consistently for this version of the VIMMP ontology documentation. To the knowledge of the authors, main developers of the EMMO at this stage include Y. Bami, J. Francisco Morgado, J. Friis, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, and D. Toti.

2. Top-level ontology and fundamental categories

2.1. European Virtual Marketplace Ontology (EVMPO)

The ontology EVMPO was developed jointly by the projects involved in establishing the EVMF (*i.e.*, VIMMP and Marketplace, with support from the EMMC-CSA project) as a common point of departure for the standardization of service-oriented semantics relevant to digital marketplace platforms in materials modelling.[2] By defining eleven fundamental paradigmatic categories, which correspond to irreducible terms that are constitutive to the paradigm underlying materials modelling marketplaces, the EVMPO provides a basic structure for the development of marketplace-level domain ontologies. These fundamental paradigmatic categories were agreed between the involved projects as follows:

1. **evmpo:assessment**, *i.e.*, a proposition on the accuracy or performance of an entity or an a expression of trust in an entity. Corresponding domain ontology: VIVO, *cf.* Section 3.7.
2. **evmpo:calendar_event**, *i.e.*, a meeting or activity which is scheduled or can be scheduled; this is defined to be equivalent with **Vevent** from the W3C iCalendar ontology (iCal) with time zones as datatypes, *cf.* Connolly and Miller [Con05]. Corresponding domain ontology: OTRAS, *cf.* Section 3.4.
3. **evmpo:communication**, *i.e.*, any message (or an attachment or part of a message) that is communicated. Corresponding domain ontology: VICO, *cf.* Section 3.5.
4. **evmpo:information_content_entity**; *e.g.*, a journal article, a data set, or a graph. This concept is defined to be equivalent with **IAO_0000030**, labelled “information content entity” through **rdfs:label**, from the Information Artifact Ontology (IAO), *cf.* Ceusters [Ceu12]. Corresponding domain ontologies: OTRAS and VISO, *cf.* Sections 3.4 and 3.6.
5. **evmpo:infrastructure**, *i.e.*, infrastructure of an EVMF-interoperable platform (*e.g.*, related to data, hardware, and software). Corresponding domain ontologies: MACRO and VISO, *cf.* Sections 3.1 and 3.6.
6. **evmpo:interpreter**; this concept is defined to be the same as **emmo-semiotics:Interpreter** from the nominalist revision of Peirce’s semiotics, based on the semiotic triad sign – object – interpretant as included in the EMMO [Fra20]; therein, for any given triad, the interpreter is the entity that carries out the semiosis, taking the sign (a representamen) as an input and producing the interpretant (another representamen) as an output. Therefore, any potential agent or communicating entity at EVMF-interoperable infrastructures is an interpreter. Corresponding domain ontology: VICO, *cf.* Section 3.5.
7. **evmpo:material**, *i.e.*, an amount of a physical substance (or mixture of substances) that is part of a more comprehensive real-world object; this concept is defined to be the same as **emmo-physicalistic:Material** from the EMMO [Fra20]. Corresponding domain ontologies: OSMO and VIVO, *cf.* Sections 3.3 and 3.7.

[2] Documented version: EVMPO v1.3.1, dated 27th June 2020. Significant contributions to EVMPO development by Y. Bami, W. L. Cavalcanti, E. Ghedini, A. Hashibon, and G. J. Schmitz are acknowledged.

8. evmpo:model, *i.e.*, an entity that represents a physical object or process by direct similitude and/or within a mathematical framework; this concept is defined to be the same as emmo-models:Model from the EMMO [Fra20]. Corresponding domain ontologies: OSMO, VISO, and VOV, *cf.* Sections 3.3, 3.6, and 3.8.
9. evmpo:process, *i.e.*, the temporal evolution of one or multiple entities. Corresponding domain ontologies: MMTO, OSMO, and VISO, *cf.* Sections 3.2, 3.3, and 3.6.
10. evmpo:product, *i.e.*, a good or service – which can be offered either on a EVMF-interoperable digital marketplace or off-site. Corresponding domain ontologies: MACRO, MMTO, and OTRAS, *cf.* Sections 3.1, 3.2, and 3.4.
11. evmpo:property, *i.e.*, an entity that is determined by an observation process, involving a specific observer that perceives or measures it; this concept is defined to be the same as emmo-properties:Property from the EMMO [Fra20]. Corresponding domain ontologies: VIVO and VOV, *cf.* Sections 3.7 and 3.8.

These categories need not be disjoint; *e.g.*, evmpo:material and evmpo:product overlap, since a material can be manufactured with the intent of selling it as a commodity, by which it becomes a good. The common superclass of the fundamental paradigmatic categories is evmpo:paradigmatic_entity. Below the fundamental level, the EVMPO also includes non-fundamental entites as subclasses; *e.g.*, evmpo:simulation as a subclass of evmpo:process, and evmpo:service as a subclass of evmpo:product. Terms which are not closely related to the materials modelling marketplace paradigm itself, but may occur within a related knowledge base, are defined to be non-paradigmatic. For this purpose, the EVMPO includes evmpo:annotation as a twelvth fundamental category, which is non-paradigmatic. (The EVMPO top relation, parent to both evmpo:paradigmatic_entity and evmpo:annotation, is called evmpo:marketplace_related_entity.) The relation evmpo:has_annotation can connect any marketplace-related entity to an annotation. Below, twelve subproperties are defined, corresponding to the fundamental categories; *e.g.*, evmpo:has_assessment_annotation for annotations of an assessment, evmpo:has_calendar_event_annotation for annotations of a calendar event, *etc.*, and evmpo:has_meta_annotation for annotations of an annotation.

Consistency with the EVMPO, and by implication consistency with the EMMO, is a requirement for all components and infrastructures that aim at interoperating within the EVMF. This design ensures that while EVMF-interoperable infrastructures need to agree on the definition of the most important entities, any platform retains the option to extend its own semantic base as required. To remain interoperable within the EVMF, any additional concepts need to be subsumed under fundamental categories from the EVMPO.

2.2. EMMO-VIMMP Integration (EVI)

The major ingredients of the EMMO approach to formalizing materials modelling [Fra20] are physicalist mereotopology following Varzi [Var96] and a nominalist reinterpretation of Peirce's semiotics [Pei91]. Therein, physicalist mereotopology primarily addresses the description of materials, which is extended by nominalist semiotics to describe modelling, simulation, and experiments. For aligning the VIMMP ontologies with the EMMO, a module with a scaled-down EMMO in TTL format is included, which is called EMMO1s (*i.e.*, EMMO version 1 simplified).[3] EMMO1s provides user-friendly IRIs for EMMO concepts, retaining the labels; *e.g.*, the IRI of the EMMO concept labelled "Semiosis" is given in the EMMO as emmo-semiotics:EMMO_008fd3b2_4013_451f_8827_52bceab11841, for which EMMO1s specifies the alias emmo1s:Semiosis as follows:

```
emmo1s:Semiosis a owl:Class;                                            (2.2.1)
   rdfs:label "Semiosis"^^xs:string;
   owl:sameAs emmo-semiotics:EMMO_008fd3b2_4013_451f_8827_52bceab11841.
```

For the purpose of the present documentation, in the interest of notational clarity, these concepts are referred to by the EMMO-based prefix, followed by the label (not the IRI), *e.g.*, emmo-semiotics:Semiosis. Technically, however, this is implemented in terms of the EMMO1s classes (e.g., emmo1s:Semiosis) the identity of which with the full, non-human-readable IRIs from the EMMO is established by statements in EMMO1s as above.

Fig. 2.1 shows how the EVMPO fundamental categories, *cf.* Section 2.1, are aligned with EMMO concepts through a module for EMMO-VIMMP Integration[4] (EVI). Beside the straightforward cases listed in Section 2.1 where EVMPO categories were designed to directly match EMMO concepts (*e.g.*, evmpo:material, defined by identity with emmo-physicalistic:Material), the EVI module states by subsumption (rdfs:subClassOf) that

$$\begin{aligned}
\text{evmpo:assessment} &\sqsubseteq \text{emmo-semiotics:Sign}, \\
\text{evmpo:communication} &\sqsubseteq \text{emmo-perceptual:Symbolic}, \\
\text{evmpo:information\_content\_entity} &\sqsubseteq \text{emmo-perceptual:Symbolic}, \\
\text{evmpo:infrastructure} &\sqsubseteq \text{emmo-manufacturing:Engineered}, \\
\text{evmpo:process} &\sqsubseteq \text{emmo-models:Model} \\
&\qquad \sqcup\ \text{emmo-holistic:Process}, \\
\text{evmpo:product} &\sqsubseteq \text{emmo-manufacturing:Engineered} \\
&\qquad \sqcup\ \text{emmo-semiotics:Sign},
\end{aligned} \tag{2.2.2}$$

where $\sqcup$ denotes owl:unionOf.

Qualified subsumptions such as "an annotation is a symbolic entity that is a proper part of a sign" and[5] "a business process is a model for a (physical) process," *etc.*, are stated as

$$\begin{aligned}
\text{evmpo:annotation} &\sqsubseteq \text{emmo-perceptual:Symbolic} \\
&\qquad \sqcap\ \exists \text{P}.\text{emmo-semiotics:Sign}, \\
\text{evmpo:business\_process} &\sqsubseteq \text{emmo-models:Model} \\
&\qquad \sqcap\ \exists(\text{emmo-models:hasModel})^{-1} \\
&\qquad\qquad .\text{emmo-holistic:Process},
\end{aligned} \tag{2.2.3}$$

[3] Documented version: EMMO1s v1.1.1, dated 20th November 2020.

[4] Documented version: EVI v1.1.9, dated 4th November 2020.

[5] Therein, evmpo:physical_process $\equiv$ emmo-holistic:Process.

$$
\begin{aligned}
\mathrm{evmpo{:}calendar\_event} \quad &\sqsubseteq \quad \mathrm{emmo\text{-}semiotics{:}Sign} \\
&\qquad \sqcap\ \exists \mathrm{S}.\mathrm{emmo\text{-}holistic{:}Process}, \\
\mathrm{evmpo{:}material\_property} \quad &\sqsubseteq \quad \mathrm{emmo\text{-}properties{:}Property} \\
&\qquad \sqcap\ \exists (\mathrm{emmo\text{-}properties{:}hasProperty})^{-1} \\
&\qquad\qquad .\mathrm{emmo\text{-}physicalistic{:}Material}, \\
\mathrm{evmpo{:}service} \quad &\sqsubseteq \quad \mathrm{emmo\text{-}semiotics{:}Sign} \\
&\qquad \sqcap\ \exists \mathrm{S}.\mathrm{emmo\text{-}holistic{:}Process}.
\end{aligned}
$$

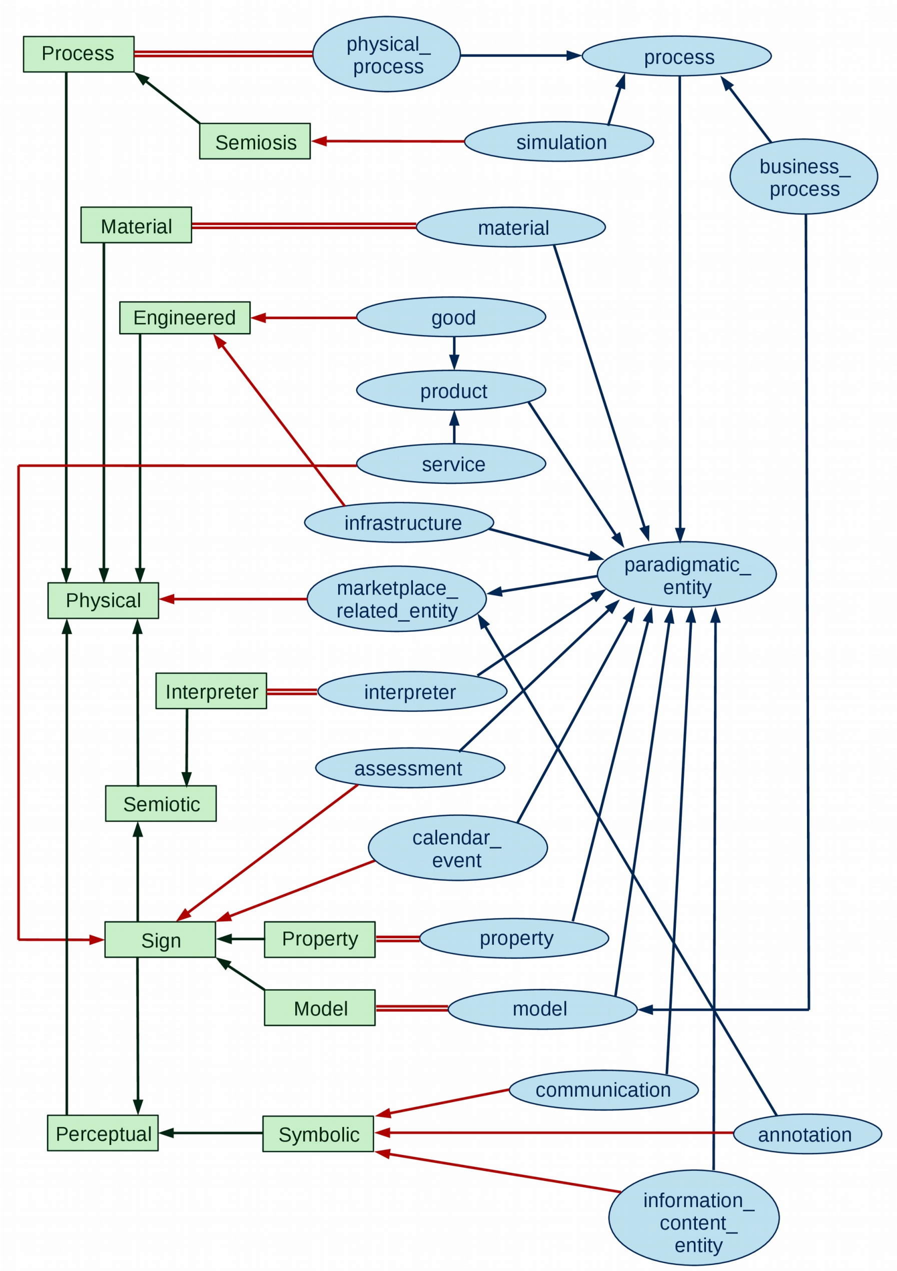


Figure 2.1. Fundamental categories, superclasses, and selected subclasses from the EVMPO (ellipses) together with related concepts from EMMO version 1.0.0β (rectangles); arrows between concepts denote subsumption, and double lines between concepts denote identity.

In Eq. (2.2.3), the operator ⊓ represents owl:intersectionOf, the relation P denotes proper parthood, the relation S denotes signification

$$
\begin{aligned}
\mathrm{P} \quad &\equiv \quad (\text{emmo-mereotopology:hasProperPart})^{-1}, \qquad (2.2.4)\\
\mathrm{S} \quad &\equiv \quad (\text{emmo-semiotics:hasSign})^{-1},
\end{aligned}
$$

Q^{-1} stands for [owl:inverseOf Q], and OWL DL notation is used [Baa17]. While some of the relations (*i.e.*, object properties) from the VIMMP ontologies can be immediately subsumed under EMMO relations, which is done in EVI, others require the concatenated mereosemiotic relations and modal logic from VIPRS, *cf.* Section 2.3.

2.3. VIMMP Primitives (VIPRS)

2.3.1. Datatype properties

The EMMO relations, like the concept definitions, are rooted in mereotopology and nominalist semiotics [Fra20], *cf.* Section 2.2. The VIMMP Primitives (VIPRS) module[6] extends the EMMO-based categorization of relations by three features: 1) Top-level datatype properties; 2) concatenation of mereotopological and semiotic relations, yielding mereosemiotic relations; 3) modal logic and modal squares of opposition. As summarized below, this significantly amplifies the ways in which the EMMO-based top-level semantic interoperability architecture can be applied to the relations characterizing metadata from the VIMMP marketplace-level domain ontologies.

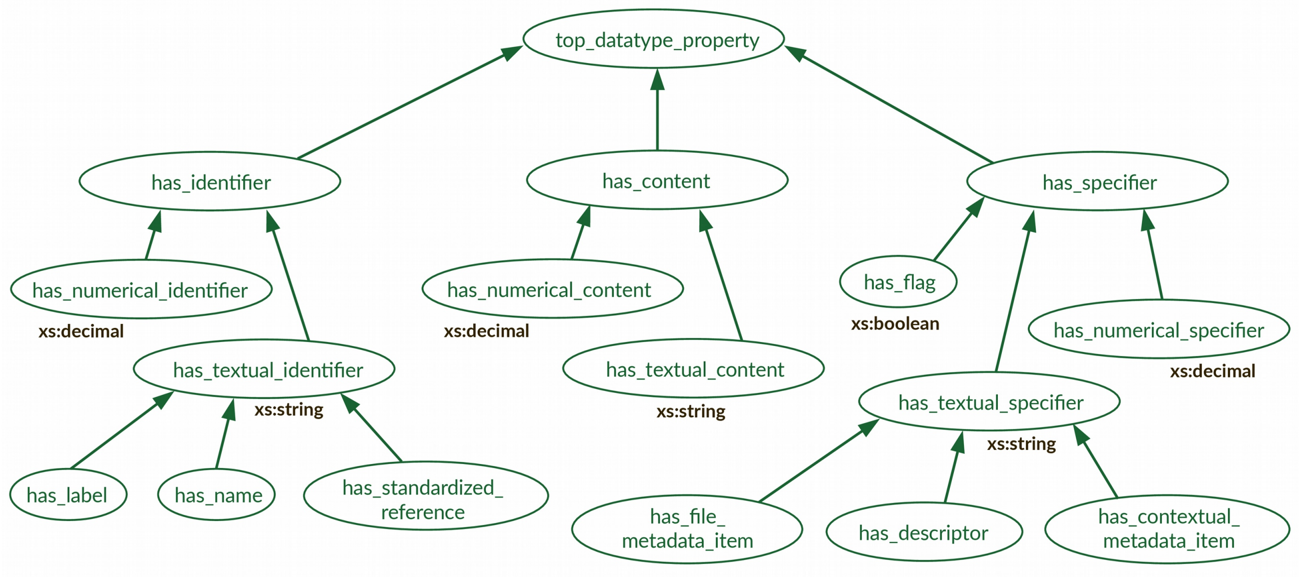


Figure 2.2. Hierarchy of datatype properties from VIPRS; arrows denote subsumption.

With minor exceptions, datatype properties (owl:DatatypeProperty) are absent from the EMMO; by the domain ontologies, however, datatype properties are amply employed to associate objects with textual (xs:string), numerical (xs:decimal) and boolean (xs:boolean) attributes. At the top level, VIPRS categorizes datatype properties according to their role:

- Identification of an object is positioned below viprs:has_identifier; examples include otras:has_topic_code, which maps a materials modelling topic

[6] Documented version: VIPRS v1.1.1, dated 20th November 2020.

(otras:mm_topic) from OTRAS to a four-digit code. Each topic code uniquely corresponds to one topic, and its purpose is identification.

- Where an elementary-datatype entry is the content (or part of the content) of an object, datatype properties below viprs:has_content are used; *e.g.*, this applies to textual or numerical content of MODA form entries (in OSMO, aspects), corresponding to osmo:has_aspect_text_content and osmo:has_aspect_text_content, *cf.* Section 3.4.
- For flags and textual or numerical descriptors, specifiers, and similar elementary metadata that provide additional, contingent information on objects, viprs:has_specifier is used; *e.g.*, otras:has_cited_video_duration_seconds points to a metadata item on the length of a video – this contributes to our knowledge about the video by specification, while it does not permit its identification. Moreover, the video duration is information *about* the video content, but it is not itself the content. Hence, otras:has_cited_video_duration_seconds $\sqsubseteq$ viprs:has_specifier.

At the second level, the datatypes are distinguished (string, decimal, or boolean). Below, at a third level, the textual datatype properties are further split into subproperties according to their function; Fig. 2.2 visualizes this hierarchy.

2.3.2. Mereosemiotic relations

To support the alignment of domain-ontology relations with EMMO relations, VIPRS introduces IRIs for relations from E^+, the free semigroup over $\mathrm{E} = \{\mathrm{P}, \mathrm{P}^{-1}, \mathrm{S}, \mathrm{S}^{-1}\}$, *cf.* Eq. (2.2.4), with the product defined by concatenation. Specifically, VIPRS contains the composite relations from $\mathrm{E} \cup \mathrm{E}^2 \cup \mathrm{E}^3$, *i.e.*, obtained from up to three signification or proper parthood relations, with up to two semiotic and up to two mereological elements (*e.g.*, $\mathrm{S}^{-1} \circ \mathrm{S}^{-1} \circ \mathrm{S}^{-1}$ contains three semiotic elements and is therefore not included in VIPRS), which are not complete (*i.e.*, relating everything to everything, except for a single “universe” entity) or redundant. The latter two provisions exclude relations containing any of the factors $\mathrm{P} \circ \mathrm{P}$ and $\mathrm{P}^{-1} \circ \mathrm{P}^{-1}$, which are redundant,[7] as well as $\mathrm{P} \circ \mathrm{P}^{-1}$, which is complete.[8] The nomenclature for the IRIs encodes “is proper part of” (P) by ip, “has proper part” (P^{-1}) by hp, “is sign for” (S) by is, and “has sign” (S^{-1}) by hs. Accordingly, *e.g.*,

$$\mathrm{viprs{:}mereosemiotics\_hp\_ip\_hs} \quad \equiv \quad \mathrm{P}^{-1} \circ \mathrm{P} \circ \mathrm{S}^{-1}, \tag{2.3.1}$$

where $\mathrm{x}\ (\mathrm{P}^{-1} \circ \mathrm{P} \circ \mathrm{S}^{-1})\ \mathrm{z}$ holds whenever there is an individual y that overlaps with x, such that $\mathrm{x}\ (\mathrm{P}^{-1} \circ \mathrm{P})\ \mathrm{y}$, and for which z is as a sign, *i.e.*, $\mathrm{y}\ \mathrm{S}^{-1}\ \mathrm{z}$. The other mereosemiotic relations from VIPRS are specified in the same way.

[7] In terms of the 4D continuum spatiotemporal entities considered within the EMMO, both P and P^{-1} are idempotent; for any $\mathrm{x}\ \mathrm{P}\ \mathrm{z}$ there is a y such that $(\mathrm{x}\ \mathrm{P}\ \mathrm{y}) \wedge (\mathrm{y}\ \mathrm{P}\ \mathrm{z})$. The EMMO explicitly permits items to be “void,” *i.e.*, not to contain any physical matter [Fra20], so that continuum nature can be assumed for EMMO spacetime even concerning properties that are subject to quantization.

[8] Following EMMO mereotopology [Fra20, Var96], there is an item u, which may be labelled “the universe” or more properly speaking “the trajectory of the universe,” of which all entities except u itself are a proper part. Hence, $(\mathrm{x}\ \mathrm{P}\ \mathrm{u}) \wedge (\mathrm{u}\ \mathrm{P}^{-1}\ \mathrm{y})$ holds for all $\mathrm{x}, \mathrm{y} \neq \mathrm{u}$, by which is $\mathrm{P} \circ \mathrm{P}^{-1}$ is complete.

2.3.3. Modal relations

While the EMMO describes materials and models as such, *i.e.*, with respect to their essence, statements on necessity and possibility anchored in modal logic are metaontological [Ghe20] from the point of view of the EMMO.[9] The present domain ontologies, however, make ample use of relations that are ultimately modal, *e.g.*, when specifying capabilities (it *is possible* that $\mathbf{x}$ will be used to do $\mathbf{y}$) or requirements (if $\mathbf{x}$ occurs, $\mathbf{y}$ also *needs to* occur). To provide a top-level structure for such relations, VIPRS includes modal squares of opposition, *cf.* Fig. 2.3, based on a plurality of modes of existence[10] by which the presence of individuals in a knowledge base can be associated with well-defined semantics, *cf.* Berto and Plebani [Ber15]. Modal operators can be given a variety of interpretations, depending on the precise use that is made of the ideas of "necessity" and "possibility" [Hut04]; VIPRS retains this ambiguity in order to remain applicable to diverse types of knowledge bases and infrastructures.

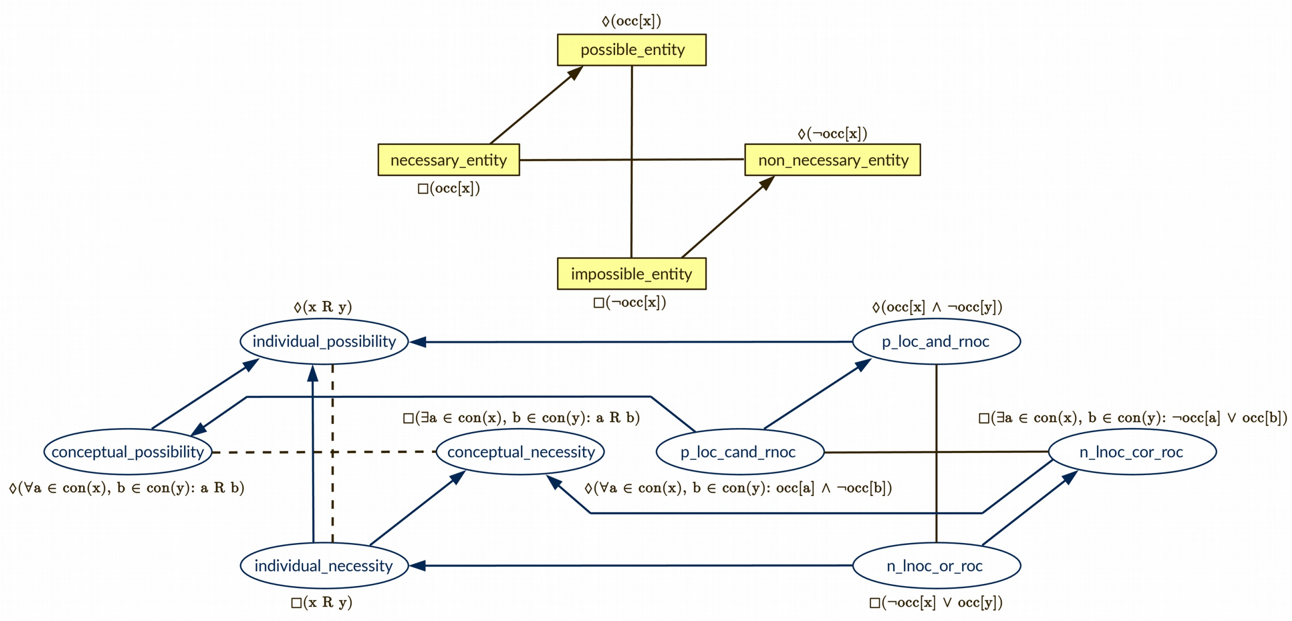


Figure 2.3. Modal squares of opposition from VIPRS. The operators $\Box$ and $\Diamond$ denote necessity and possibility, $\mathrm{occ}[\mathbf{x}]$ stands for "$\mathbf{x}$ occurs," $\mathrm{con}(\mathbf{x})$ stands for a conceptualization of $\mathbf{x}$, arrows represent subsumption (top: $\mathbf{rdfs{:}subClassOf}$, bottom: $\mathbf{rdfs{:}subPropertyOf}$), solid lines represent complementarity (top) and negation (bottom), and dashed lines connect conjugate relations.

[9] As Ghedini explains [Ghe20], following Eco [Eco75], the EMMO is a tool for formulating "the perfect lie," while statements like "this is a lie" or "this can possibly occur, but it will not necessarily occur" are beyond its scope. From this perspective, such statements are *metaontological* – beyond the EMMO.

[10] VIPRS employs the term "to occur" as in $\Diamond(\mathrm{occ}[\mathbf{x}])$, "$\mathbf{x}$ may occur," and similar, to refer to the (possible or necessary) appearance of an individual $\mathbf{x}$ in a certain type of environment, *e.g.*, as an element of a valid simulation workflow. On this basis, relations concerning the possible or necessary co-occurrence of multiple individuals are defined, *cf.* Fig. 2.3. Thereby, "occurrence" (by appearing in a certain type of environment) is not the same as "existence," *i.e.*, presence in a knowledge base; *n.b.*, for users of the VIMMP ontologies it is not necessary to familiarize themselves with these aspects – modal logic is introduced here to complement the EMMO top-level structure in a way that permits covering all relations from the present ontologies by subsumption.

3. Domain ontologies from VIMMP

3.1. Marketplace-Accessible Computational Resource Ontology (MACRO)

The ontology MACRO deals with data and hardware related resources and infrastructures [Hor20a]. In particular, MACRO contains classes and individuals representing file formats expected to occur on the VIMMP marketplace platform,[11] many of which are obtained by connecting to the EDAM ontology [Iso13]. High-level concepts from MACRO and their relation to EVMPO concepts (agent, annotation, infrastructure, service) are shown in Fig. 3.1. Complementing MACRO, the PaaSPort ontology [Bas18] can be used to describe platforms as a service (PaaS).

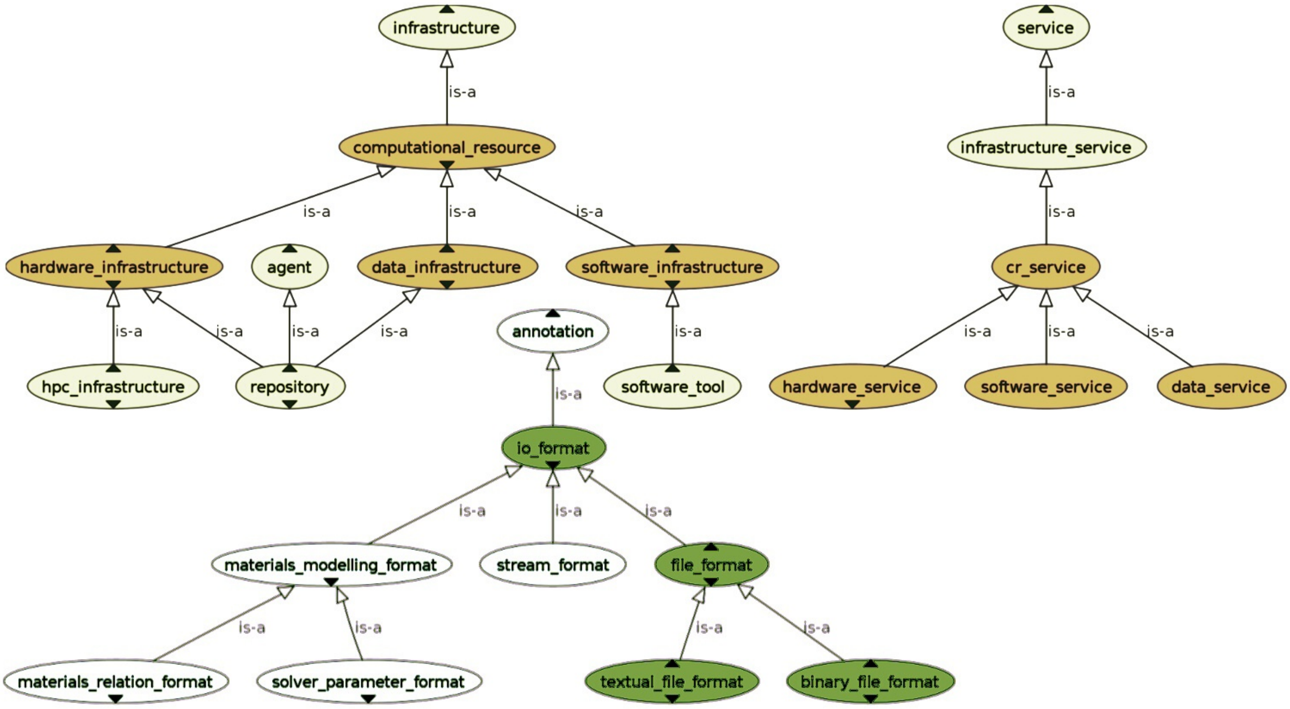


Figure 3.1. High-level part of the MACRO class hierarchy. The OWLViz protégé plugin was used to generate the diagram; arrows labelled “is-a” denote subsumption ($\sqsubseteq$), *i.e.*, $\mathrm{rdfs{:}subClassOf}$.

Selected concept definitions from MACRO:

- $\mathbf{macro{:}channel}$, *i.e.*, a data infrastructure which, in its evolution as a process, contains communication events (semioses). EMMO alignment:
$$\begin{aligned}\mathbf{macro{:}channel} \sqsubseteq \mathbf{emmo\text{-}holistic{:}Process} \sqcap \exists(\mathbf{P}^{-1}).\mathbf{emmo\text{-}semiotics{:}Semioisis}\\ \sqcap\ \mathbf{emmo\text{-}manufacturing{:}Engineered}.\end{aligned}$$
- $\mathbf{macro{:}computational\_resource}$, *i.e.*, an infrastructure that can be accessed by means of data, hardware, or software related services. EMMO alignment:
$$\mathbf{vivo{:}computational\_resource} \sqsubseteq \mathbf{emmo\text{-}manufacturing{:}Engineered}.$$

[11] Documented version: MACRO v1.1.7, dated 20th November 2020. Significant contributions to MACRO development by E. Fayolle, Y. Fournier, J.-P. Minier, P. Noyret, and V. Stobiac, the authors of VIMMP project deliverable 2.1, as well as A. Fiseni and H. Krieg are acknowledged.

- macro:infrastructure_service, *i.e.*, a service that provides access to an infrastructure. EMMO-VIPRS alignment:
  $$\begin{aligned}&\texttt{macro:infrastructure\_service}\\ &\qquad\sqsubseteq \texttt{emmo-semiotics:Sign} \sqcap \exists(\mathrm{S}).\texttt{emmo-holistic:Process}\\ &\qquad\qquad \sqcap \exists(\mathrm{S} \circ \mathrm{P}^{-1}).\texttt{emmo-manufacturing:Engineered}.\end{aligned}$$
- macro:io_format, *i.e.*, a syntactical convention to which a technical I/O implementation can adhere. EMMO alignment:
  $$\begin{aligned}&\texttt{macro:io\_format}\\ &\qquad\sqsubseteq \texttt{emmo-perceptual:Symbolic} \sqcap \texttt{emmo-semiotics:Conventional}\\ &\qquad\qquad \sqcap \texttt{emmo-perceptual:Language} \sqcap \exists\mathrm{P}.\texttt{emmo-semiotics:Sign}.\end{aligned}$$
- macro:model_database, *i.e.*, a repository that can act as a model provider. EMMO alignment: $\texttt{macro:model\_database} \sqsubseteq \texttt{emmo-manufacturing:Engineered} \sqcap \texttt{emmo-semiotics:Interpreter}$.

Selected relations (object properties) from MACRO:

- macro:has_channel_member; points to an agent (*i.e.*, communicating entity) that participates in communicating through a channel. Domain: macro:channel; range: evmpo:agent.
  EMMO alignment: $\texttt{macro:has\_channel\_member} \sqsubseteq \texttt{emmo-holistic:hasParticipant}$.
- macro:has_granularity; points to the granularity level to which the entities represented in an I/O format belong. Domain: macro:materials_modelling_format; range: osmo:granularity_level.
  EMMO-VIPRS alignment: $\texttt{macro:has\_granularity} \sqsubseteq (\mathrm{S}^{-1} \circ \mathrm{P}^{-1}) \sqcap (\mathrm{S} \circ \mathrm{P}^{-1} \circ \mathrm{S}^{-1})$.
- macro:is_io_format_of; points to a software tool that can process files in a given I/O format. Domain: macro:io_format; range: viso:software_tool.
  EMMO-VIPRS alignment: $\texttt{macro:is\_io\_format\_of} \sqsubseteq \mathrm{S} \circ \texttt{viprs:can\_cooccur\_with}$.
- macro:provides_access_to; points to a service that can be accessed through the given infrastructure. Domain: macro:infrastructure; range: macro:infrastructure_service. EMMO-VIPRS alignment:
  $$\texttt{macro:provides\_access\_to} \sqsubseteq (\mathrm{S} \circ \mathrm{P}^{-1}) \sqcap \texttt{viprs:satisfies\_requirement\_of}.$$

3.2. Materials Modelling Translation Ontology (MMTO)

3.2.1. Business case (BC), industrial case (IC), and translation case (TC)

The ontology MMTO deals with the paradigm of materials modelling translation, *i.e.*, translation from engineering practice to modelling and simulation, and from the simulation outcome back to an actionable decision.[12] The role of the materials modelling translator is defined in the EMMC Translators' Guide (ETG), *cf.* Hristova *et al.* [Hri19]; a translator needs to be able to bridge the "language gap" between industrial end users as well as academic model providers and software owners. The work of a translator aims at delivering not only

[12] Documented version: MMTO v1.3.6, dated 20th November 2020. Significant contributions to MMTO development by P. Klein, N. A. Konchakova, and B. Schembera are acknowledged.

modelling results but also a valuable and beneficial solution for a problem from industrial engineering practice. An instance of the materials modelling translation process, some agreed features of which are codified by the ETG and the EMMC Translation Case Template (ETCT) [Emm17], is referred to as a translation case (TC). According to these specifications, a materials modelling translation project begins with exploring and understanding the business case (BC) and the industrial case (IC), or multiple relevant BCs and/or ICs, which characterize socioeconomic objectives and boundary conditions.

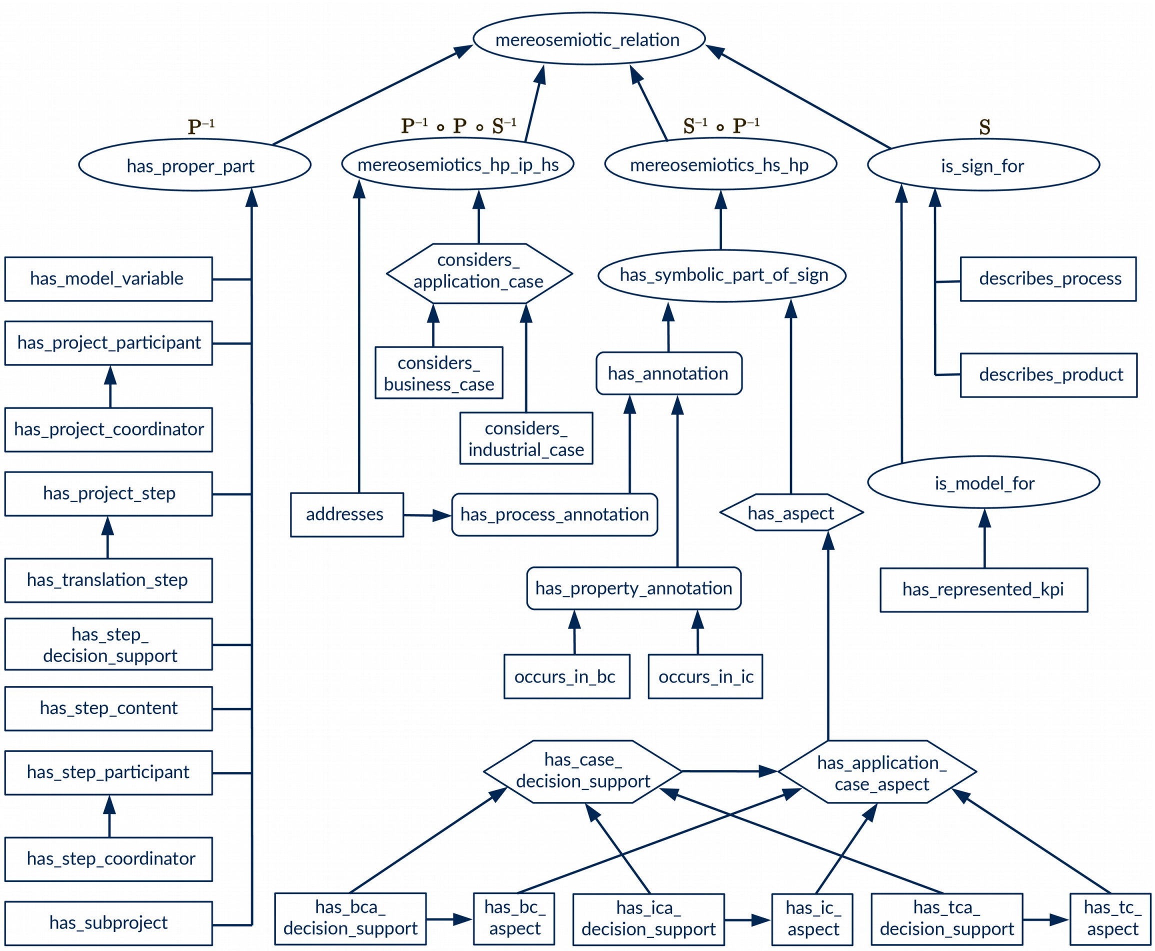


Figure 3.2. Hierarchy of object properties, where arrows represent the transitive reduction of rdfs:subPropertyOf, for relations defined in the MMTO (rectangles), showing their subsumption under relations from OSMO (hexagons), the EVMPO (rounded boxes), and VIPRS (ellipses).

In MODA graphs [Cen18], there are four types of vertices, here referred to as *sections*:

1. Use case, *i.e.*, the physical system to be simulated, including information on the given and desired physical properties.
2. Model (osmo:materials_model), specified by the underlying *governing equations* (GEs), *i.e.*, one or multiple *physical equations* (PEs) and *materials relations* (MRs).
3. Solver, *i.e.*, the numerical solution of the model – defined with a strict limitation to considering exactly the variables that occur in the GEs explicitly (and nothing else).
4. Processor, *i.e.*, any computational operation beyond the above; in particular, this includes and processing activity done by a simulation code that goes beyond the immediate solution of underlying governing equations (*e.g.*, aggregated output).

Additionally, the MODA forms include a cover sheet that is similarly structured; this part of a MODA provenance description here corresponds to osmo:simulation_overview, which is also a subclass of osmo:section. For each section, the MODA standard contains a list of text fields, which are here referred to as *aspects*, through which detailed information can be provided; however, since this is plain text, it is usually not immediately possible to extract semantically annotated content from this representation automatically. Since it is given as an ontology, aspects from the MMTO (and from OSMO, *cf*. Section 3.3), corresponding to plain-text form entries in MODA, can contain links to entities defined elsewhere in the semantic web which can be immediately processed computationally, and to which automated reasoning can be applied. In this way, *e.g.*, a use case becomes an osmo:use_case entity. In MODA, a section can only be described in terms of (textual and numerical) elementary data; by using the relation osmo:has_aspect_object_content, *cf*. Section 3.3.3, it becomes possible to point to content provided anywhere on the semantic web, including individuals and classes from the VIMMP marketplace-level domain ontologies. The MMTO generalizes this approach from MODA to also cover the translation-related concepts from the ETCT and the ETG [Emm17, Hri19]: Universals for BCs (mmto:business_case), ICs (mmto:industrial_case), and TCs (mmto:translation_case) are defined to be subclasses of osmo:application_case, by which they can be dealt with in a similar way as the sections from OSMO. The subsumption of relations from the MMTO under EMMO relations and composite mereosemiotic relations from VIPRS (via OSMO and EVMPO) is visualized in Fig. 3.2, and the class hierarchy of the section branch of the MMTO and OSMO is visualized in Fig. 3.3.

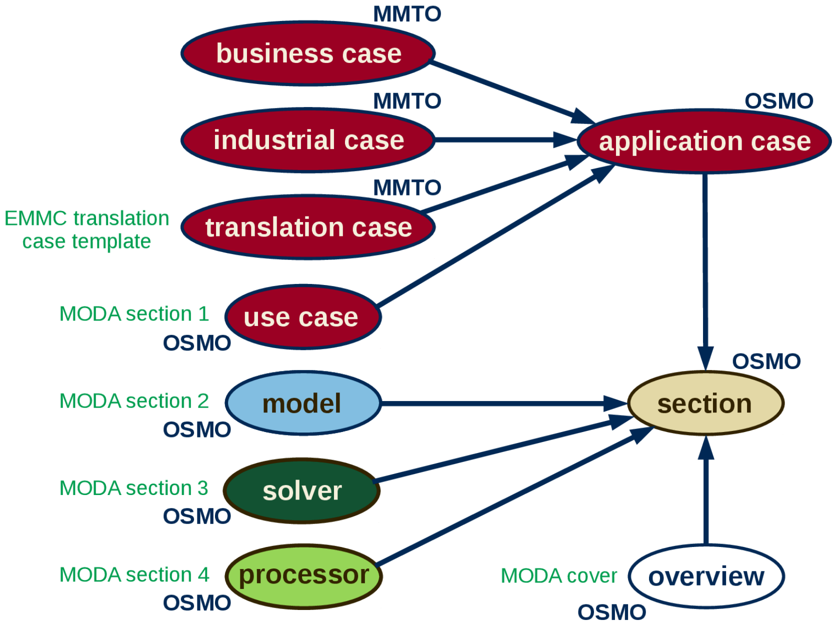


Figure 3.3. Section branch of MMTO and OSMO, *cf*. Section 3.3; arrows denote subsumption ($\sqsubseteq$).

The TC aspects directly correspond to the ETCT text fields [Emm17], except that the MMTO permits the provision of semantically characterized content. A business case can represent any purely economic consideration or an optimization problem at the management level, whereas an IC refers to an industrial engineering problem or an optimization problem at the technical or research and development level. Within the translation process, a suitable approach based on modelling and simulation is identified and carried out; subsequently, the outcome is translated back to support an actionable decision at the BC and IC levels. Thus, the MMTO is also a tool for representing exchange of information during translation processes (*e.g.*, employing KPIs as logical variables), which may be represented by a workflow analogous to MODA and the enhanced LDT workflow graphs described in Section 3.3.2. An possible exchange of communications taking place during a translation process (ordered as a sequence in time from top to bottom) is depicted in Fig. 3.4 together with the class hierarchy of the relevant branch of the MMTO.

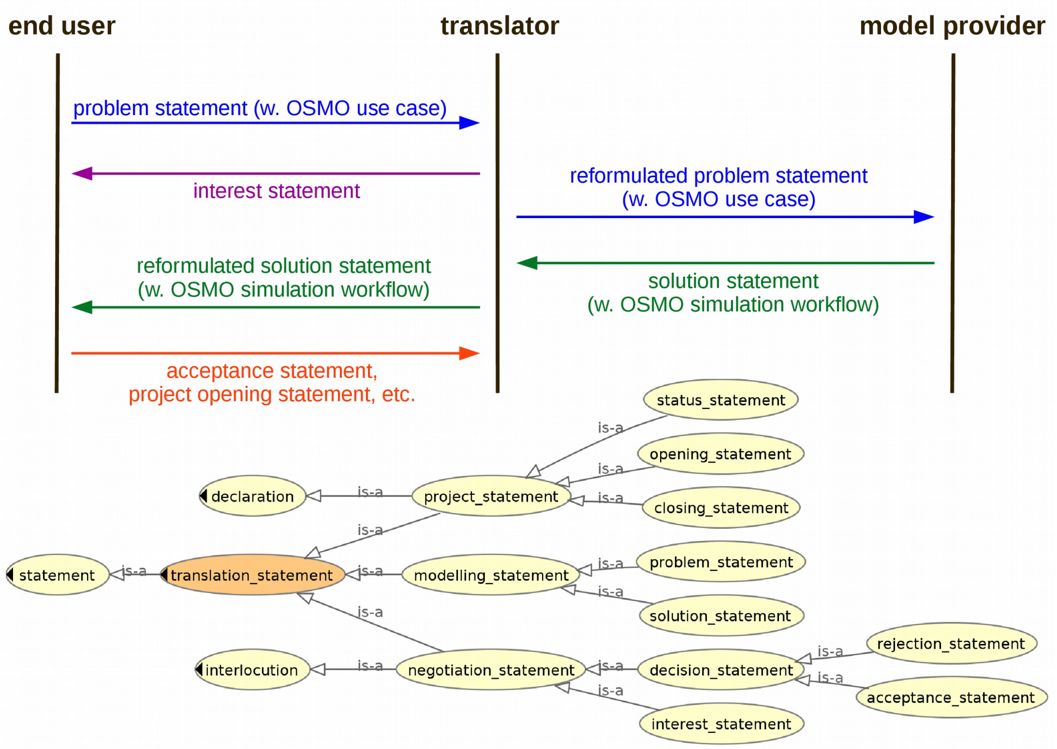


Figure 3.4. Top: Possible sequence of messages exchanged at a digital marketplace during negotiation of a materials modelling translation project. Bottom: Class hierarchy for MMTO concepts related to exchange between potential partners during such interactions; the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption (⊑), *i.e.*, rdfs:subClassOf.

3.2.2. Key performance indicators (KPIs)

In business administration and management, a KPI is understood to be a descriptor (indicator) underlying process and product opimization and ultimately characterizing some feature or property that can serve as a selling argument. The orientation toward marketing reflects a point of view corresponding to organizational roles that are comparably distant from research and development, *e.g.*, in sales or high-level management. In scenarios that arise in the context of such organizational roles, it necessarily appears to be most crucial to address concerns that are immediately relevant to business-to-administration (B2A), business-to-business (B2B), and business-to-customer (B2C) relations [Bar12]. In the MMTO, the concept mmto:key_performance_indicator is reserved for scalar quantities that are relevant for characterizing, modelling, or optimizing such scenarios. On this basis, from the point of view of a materials modelling translator, two major distinctions need to be made:

1. Some KPIs are closely related to human sentience (aesthetics, haptics, taste, *etc.*). Studies aiming at gaining information on these quantities typically rely on market research and other empirical methods that involve human subjects; such indicators are referred to as subjective KPIs (mmto:subjective_kpi). Obversely, an objective KPI (mmto:objective_kpi) can be determined by a standardized process, *e.g.*, a measurement, experiment, or simulation, the result of which (assuming that it is conducted correctly) does not depend on the person that carries it out.
2. An objective KPI is technological (mmto:technological_kpi) if it is observed or measured within a technical or experimental process, referring directly to properties of the real product or manufacturing process; properties of a model, which are determined by simulation, are computational KPIs (mmto:computational_kpi).

The distinction between subjective and objective KPIs is similar to that between critical-to-customer (CTC) and critical-to-quality (CTQ) measures [Emm18, Mac18, Pri13]. The formulation given above, however, is more closely related to concepts from the EMMO. Due to the underlying approach to semiotics [Fra20, Pei91], it is straightforward in the EMMO to categorize signs by the way in which their interpretation depends on the subjective impression of an interpreter or observer: In particular, the same distinction between subjective properties (emmo-properties:SubjectiveProperty) and objective properties (emmo-properties:ObjectiveProperty) is made in the EMMO; accordingly, the present approach supports a straightforward alignment of the MMTO with the EMMO and the approach to interoperability guided by the EMMC and implemented within the EVMF.

3.3. Ontology for Simulation, Modelling, and Optimization (OSMO)

3.3.1. OSMO – the ontology version of MODA

The ontology for simulation, modelling, and optimization[13] was developed as the ontology version of MODA, making workflow representations machine processable, semantically interoperable with community platforms, and amenable to automated reasoning [Hor20b]. Where a physically based modelling approach is followed, *physical equations* are employed jointly with *materials relations* that parameterize and complement the physical equations, *e.g.*, to describe a particular substance. The combination of PEs and MRs is referred to as the system of governing equations, *cf.* Section 3.2.1; on the basis of RoMM [Deb17], common PE types are subdivided into four groups according to their granularity level: Electronic, atomistic, mesoscopic, and continuum, *cf.* Fig. 3.5.

The detailed description of section individuals (*cf.* Fig. 3.3) in OSMO by section aspects and their textual, numerical, or object content is closely aligned with the corresponding

[13] Documented version: OSMO v1.7.5, dated 19th November 2020. Significant contributions to OSMO development by G. Boccardo, P. Carbone, W. L. Cavalcanti, M. Chiricotto, J. D. Elliott, A. Fiseni, H. Krieg, V. Lobaskin, P. Neumann, C. Niethammer, P. Schiffels, and J. Vrabec are acknowledged.

text-field specifications from MODA [Cen18], *cf.* Fig. 3.6. By providing a common semantic basis for workflows that were designed with different tools, OSMO can be employed to consistently integrate data provenance descriptions for materials modelling data from diverse sources [Hor20b].

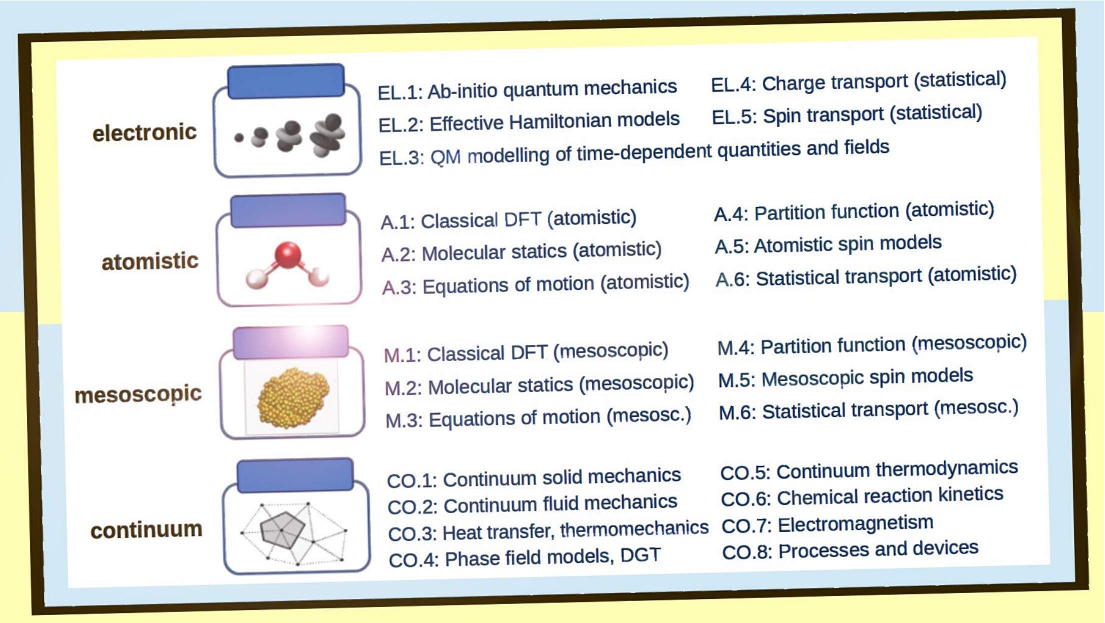

electronic
EL.1: Ab-initio quantum mechanics
EL.2: Effective Hamiltonian models
EL.3: QM modelling of time-dependent quantities and fields
EL.4: Charge transport (statistical)
EL.5: Spin transport (statistical)
atomistic
A.1: Classical DFT (atomistic)
A.2: Molecular statics (atomistic)
A.3: Equations of motion (atomistic)
A.4: Partition function (atomistic)
A.5: Atomistic spin models
A.6: Statistical transport (atomistic)
mesoscopic
M.1: Classical DFT (mesoscopic)
M.2: Molecular statics (mesoscopic)
M.3: Equations of motion (mesosc.)
M.4: Partition function (mesoscopic)
M.5: Mesoscopic spin models
M.6: Statistical transport (mesosc.)
continuum
CO.1: Continuum solid mechanics
CO.2: Continuum fluid mechanics
CO.3: Heat transfer, thermomechanics
CO.4: Phase field models, DGT
CO.5: Continuum thermodynamics
CO.6: Chemical reaction kinetics
CO.7: Electromagnetism
CO.8: Processes and devices

Figure 3.5. Taxonomy of physical equation types (subclasses of osmo:physical_equation_type) implemented on the basis of the categorization in the Review of Materials Modelling [Deb17].

3.3.2. Visualization of logical data transfer (LDT)

Logical data transfer (LDT) notation [Hor20b] clarifies how the use case, model, solver, and processor entities relate to each other in a MODA workflow representation [Cen18]. In LDT notation, *cf.* Fig. 3.7, ellipses represent use cases, models, solvers, and processors (*i.e.*, sections); green circles and green arrows represent coupling and linking of elements, dependencies concerning the order of execution, and aspects related to concurrency and synchronization. Blue arrows point from use cases and models to the part of the workflow to which these elements apply. Triangles are logical resources, describing how information is transferred between the sections, pointing from the source to the destination; if a triangle is filled, this denotes that a user interaction can occur.

The visualization elements from LDT notation have a direct correspondence with concepts and relations from OSMO; *e.g.*, coupling and linking symbolized by green arrows correspond to the relations osmo:is_coupled_with and osmo:is_linked_to, and flow of information represented by lines between logical resources (triangles) and sections (ellipses) corresponds to osmo:logical_access entities that relate to a logical resource by osmo:has_resource and to a section by osmo:has_access_point, *cf.* Fig. 3.8. The LDT representation therefore corresponds to an enriched version of a MODA graph; by removing logical resources, details on iterations (represented in OSMO by relations between “virtual graphs” and “concrete graphs”), *etc.*, a conventional MODA description can be obtained. Similarly, the usual human-readable MODA forms can be obtained by reducing all OSMO aspects to an elementary numerical or textual description.

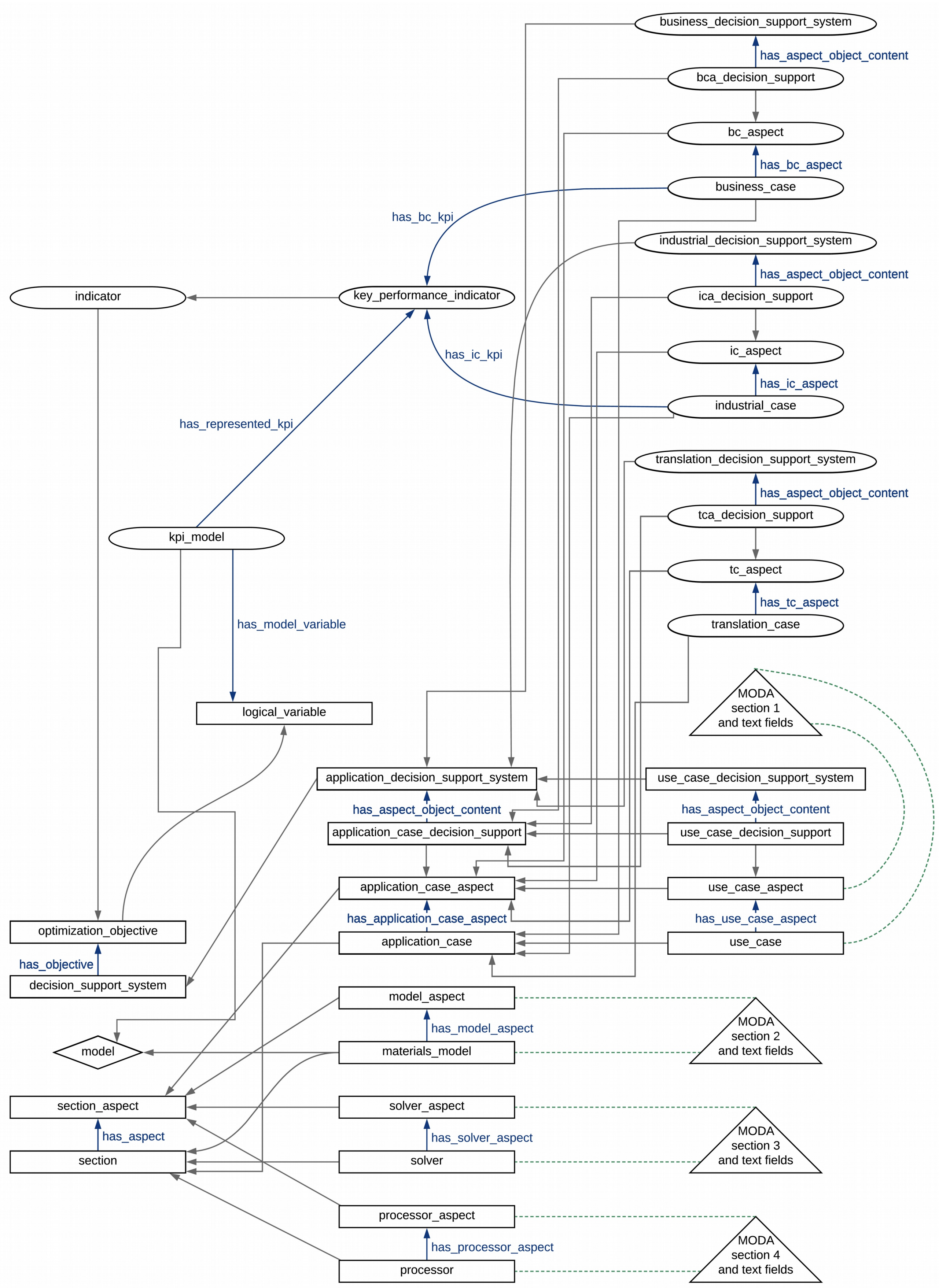


Figure 3.6. Selected concepts from OSMO (rectangles), the MMTO (rounded boxes), the EMMO and EVMPO (diamond) with selected object properties (blue arrows) and correspondences with MODA (dashed lines); grey arrows denote subsumption.

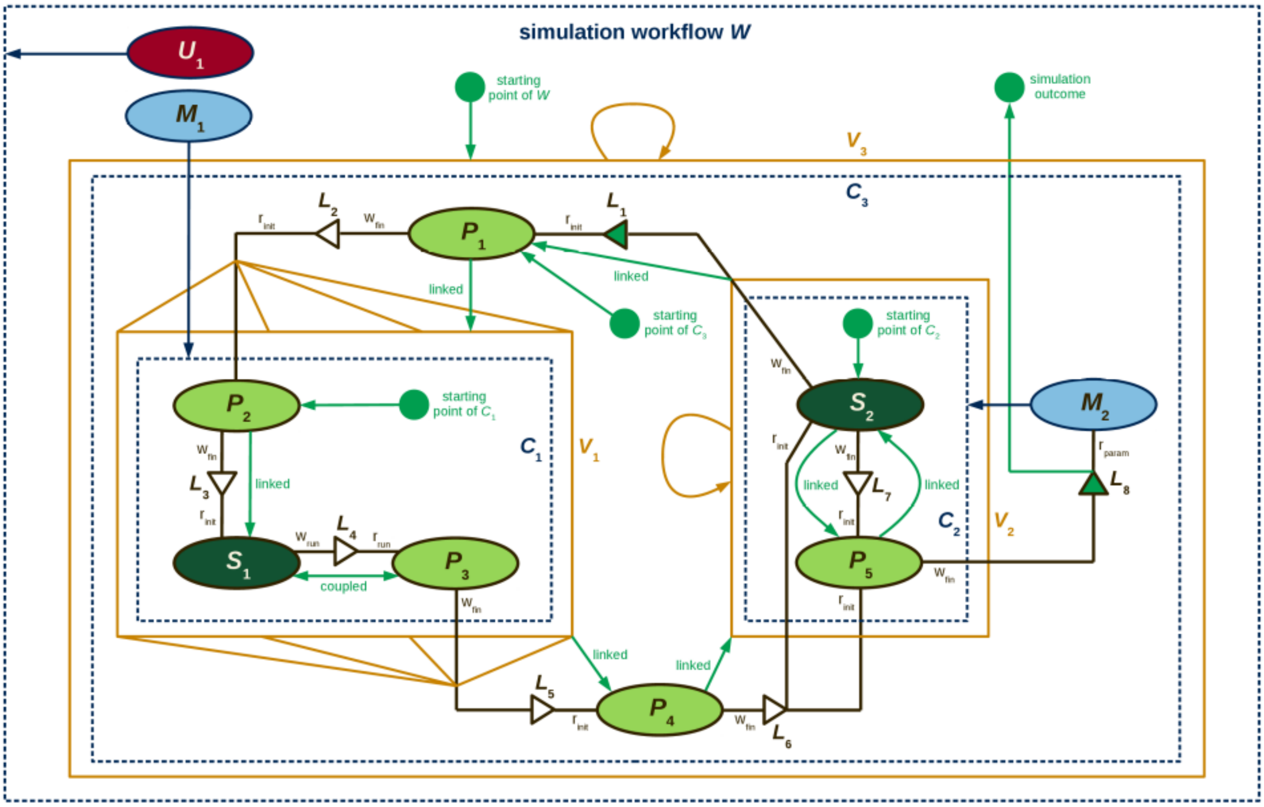


Figure 3.7. Example simulation workflow in LDT notation; scenario: Molecular-simulation based automated parameterization of a phenomenological equation of state [Hor20b, Rut15].

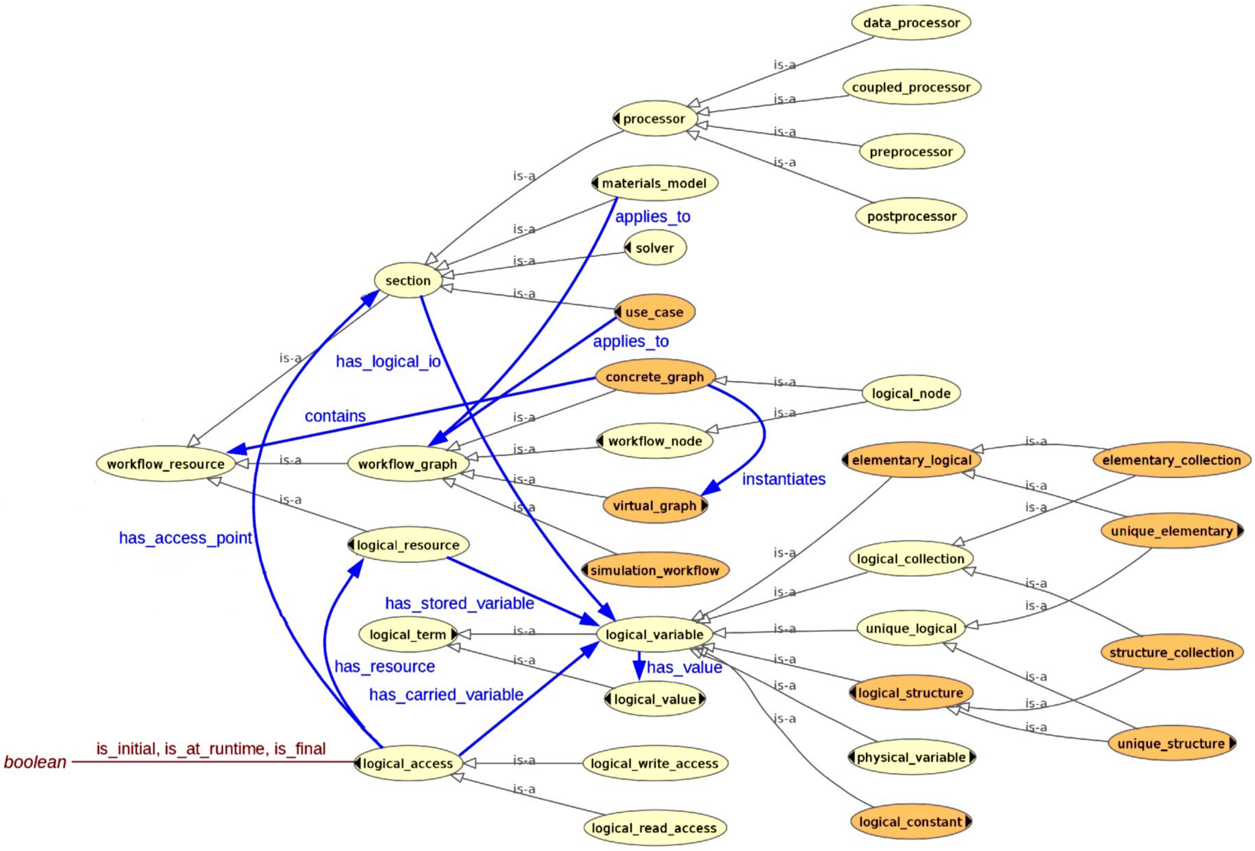


Figure 3.8. Workflow resource branch of OSMO [Hor20b]: Selected concepts (ellipses), relations (blue arrows), and datatype properties (solid line); the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption (⊑), *i.e.*, rdfs:subClassOf.

3.3.3. Illustration

Selected concept definitions from OSMO:

- osmo:condition, *i.e.*, a statement concerning values of properties and/or parameters and/or their relation to each other. Subclasses include mmto:kpi_model.
  EMMO alignment: osmo:condition $\sqsubseteq$ emmo-math:Mathematical.
- osmo:einecs_listed_material, *i.e.*, an EC listed material from the European Inventory of Existing Commercial Chemical Substances (EINECS), which can be identified by an EC number; analogous: osmo:cas_listed_material, identified by a CAS number.
  EMMO alignment: osmo:einecs_listed_material $\sqsubseteq$ emmo-physicalistic:Material.
- osmo:logical_variable, *i.e.*, a term that can be exchanged by interaction with logical resources. Subclasses, including osmo:unique_elementary (for scalar variables) and osmo:optimization_objective, are shown in Figs. 3.6 and 3.8.
  EMMO alignment: osmo:logical_variable $\equiv$ emmo-math:Variable.
- osmo:materials_relation, *i.e.*, a materials relation (MR) as defined by RoMM [Deb17], *cf.* MODA subsection 2.4 [Cen18].
  EMMO alignment: osmo:materials_relation $\equiv$ emmo-models:MaterialRelation.
- osmo:section, defined on the basis of MODA [Cen18], generalized to account for BCs, ICs, and TCs, *cf.* Section 3.2. The subclasses are shown in Figs. 3.3 and 3.6.
  EMMO alignment: osmo:section $\sqsubseteq$ emmo-physical:Physical.
- osmo:section_aspect, *i.e.*, a descriptor of an OSMO section, following the approach from MODA [Cen18]. Subclasses are shown in Fig. 3.6.
  EMMO alignment: osmo:section_aspect $\sqsubseteq$ emmo-physical:Physical.
- osmo:workflow_graph, *i.e.*, a model, a simulation workflow, or a composition of constituent elements and aspects thereof that can be represented in LDT notation. Subclasses are shown in Fig. 3.8.
  EMMO alignment: osmo:workflow_graph $\sqsubseteq$ emmo-perceptual:WellFormedSymbolic.

Selected relations (object properties) from OSMO:

- osmo:contains; (G contains R): R occurs as a proper part/component of graph G.
  Domain: osmo:concrete_graph; range: osmo:workflow_resource.
  EMMO alignment: osmo:contains $\sqsubseteq$ P^{-1}.
- osmo:has_aspect; points to an aspect associated with the respective section.
  Domain: osmo:section; range: osmo:section_aspect.
  EMMO-VIPRS alignment: osmo:has_aspect $\sqsubseteq$ viprs:has_symbolic_part_of_sign.
- osmo:has_aspect_object_content; points to an object entry associated with an aspect. Domain: osmo:section_aspect; range: evmpo:marketplace_related_entity.
  EMMO-VIPRS alignment: osmo:has_aspect_object_content $\sqsubseteq$ $P^{-1} \circ S$.
- osmo:has_logical_io; points to information required or produced by a section. Domain: osmo:section; range: osmo:logical_variable.
  EMMO-VIPRS alignment: osmo:has_logical_io $\sqsubseteq$ viprs:has_symbolic_part_of_sign.

- osmo:has_value; points to a value assigned to a logical variable. Domain: osmo:logical_variable; range: osmo:logical_value.
  EMMO-VIPRS alignment: osmo:has_logical_io ⊑ viprs:has_symbolic_part_of_sign.
- osmo:has_variable_unit; points to the unit to be associated with any assigned decimal values. Domain: osmo:elementary_logical; range: vivo:unit.
  EMMO alignment: osmo: has_variable_unit ⊑ emmo-metrology:hasReferenceUnit.
- osmo:is_linked_to; if (F is linked to G) holds, F and G cannot be executed concurrently; one side depends on the completion of the other side. Domain: osmo:workflow_graph; range: osmo:workflow_graph.
  EMMO-VIPRS alignment: osmo:is_linked_to ⊑ viprs:mutual_requirement.

3.4. Ontology for Training Services (OTRAS)

3.4.1. Training events and didactics

The ontology OTRAS can be employed to annotate any training resources in the field of materials modelling [Hor20a], *i.e.*, both training documents (such as manuals or videos) and training events (lectures, seminars, summer schools, workshops, *etc.*). In OTRAS, such resources are referred to as *carriers*.[14] For information on training courses, syllabi, *etc.*, the Course Curriculum and Syllabus Ontology (CCSO) is employed [Kat18]. Futhermore, the IAO is applied to documents, in accordance with the EVMPO. The high-level structure of OTRAS is shown in Fig. 3.9. While the CCSO covers much of the required domain at an abstract level, a dedicated standardization effort is required to characterize the semantic space with respect to training contents specifically in the field of materials modelling. For this purpose, OTRAS includes a formalism by which learning outcomes and expert competencies can be described and a taxonomy of topics in materials modelling.

Concerning didactics, the normal form of a learning-outcome specification to be used with OTRAS is given as follows:

"Upon successfully completing $\mathbf{X}_1$, participants can $\mathbf{X}_2$ (3.4.1)
with respect to $\mathbf{X}_3$ by doing $\mathbf{X}_4$; for example, $\mathbf{X}_5$."

Therein, $\mathbf{X}_1$ is the course or training material (carrier) for which a learning outcome is stated – the learning outcome is associated with a syllabus, describing the didactic approach, through the relation otras:aims_to ⊑ ccso:aimsToLO. If a competency is asserted as such, irrespective of how it has been acquired, $\mathbf{X}_1$ can be absent, in particular, wherever the relation vico:has_competency from VICO, *cf.* Section 3.5, is used to characterize the background of an evmpo:expert. The entities $\mathbf{X}_2$, $\mathbf{X}_3$, $\mathbf{X}_4$, and $\mathbf{X}_5$ are *specifiers* (otras:specifier) of the learning outcome:

[14] Documented version: OTRAS v1.1.3, dated 20th November 2020. Significant contributions to OTRAS development by B. Andreon, E. Bayro Kaiser, H. Brüning, W. L. Cavalcanti, J. Díaz Brañas, J. D. Elliott, A. Fiseni, M. Lísal, I. Pagonabarraga Mora, B. Planková, P. Schiffels, A. Scotto di Minico, and K. Šindelka are acknowledged.

- X_2 specifies the *operator* of the learning outcome (class otras:operator_specifier); a catalogue of operators with three-digit operator codes is included, *cf.* Section 3.4.3.
- X_3 specifies the *operand* of the learning outcome (class otras:operand_specifier); the operand can be formulated in terms of one or multiple topics, *cf.* Section 3.4.2.
- X_4 specifies the *implementation* (class otras:implementation_specifier), describing how the competency is carried out in practice (*e.g.*, "by writing C++ codes" or "by carrying out appropriate series of DPD simulations"); this specifier is optional.
- X_5 specifies an *example* (class otras:example_specifier), illustrating how the competency might be applied to a particular special case (*e.g.*, "if asked to develop a molecular model for caffeine, the participant might consider a rigid coarse grained model consisting of multiple Mie interaction sites"). This specifier is also optional.

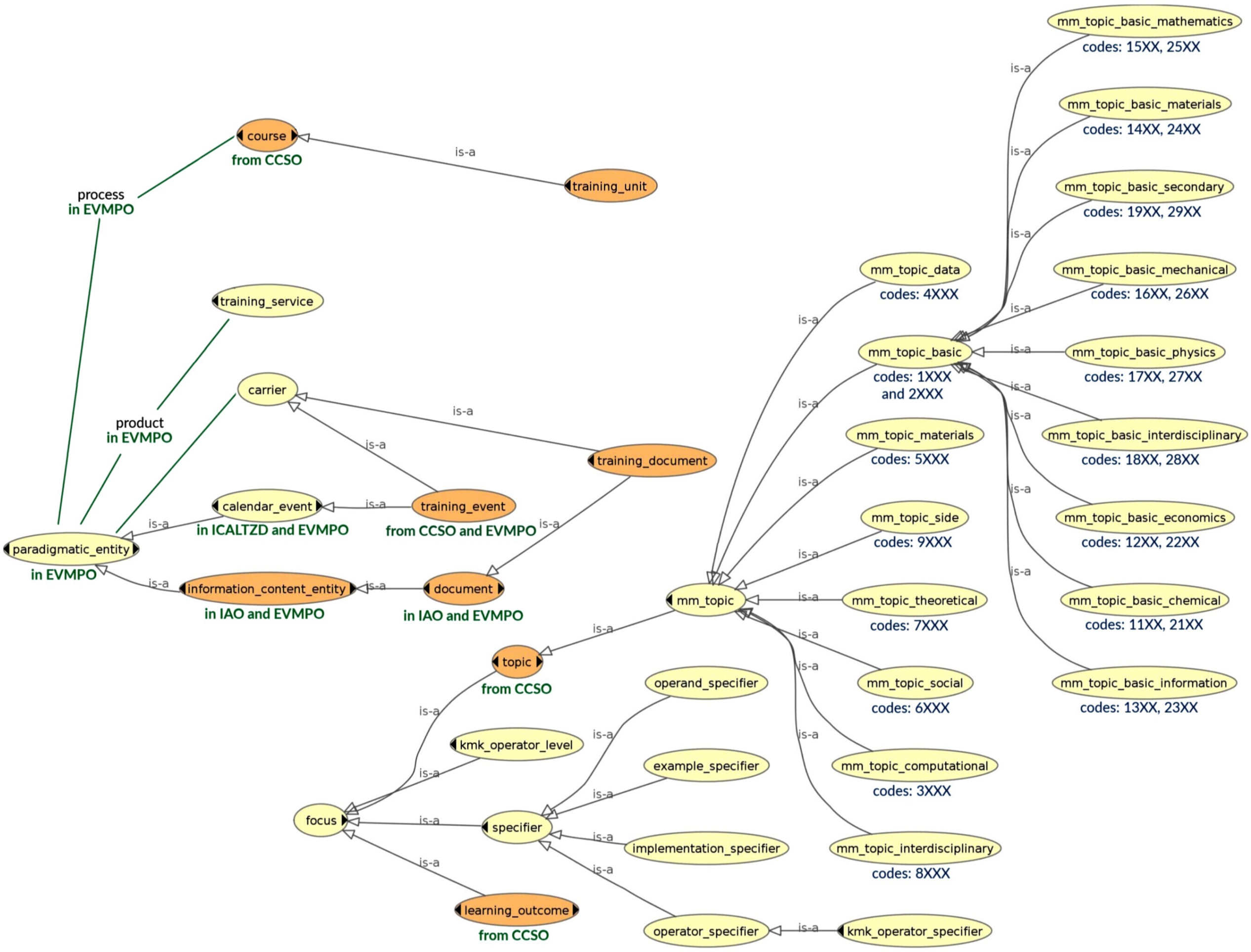


Figure 3.9. Fragment of the OTRAS class hierarchy; the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption ($\sqsubseteq$), *i.e.*, rdfs:subClassOf. Here, ICALTZD refers to the iCal ontology with time zones as datatypes [Con05].

3.4.2. Taxonomy of topics in materials modelling

At the first hierarchy level, the topics relevant to the domain of materials modelling (otras:mm_topic) are categorized as follows:

- Concept otras:mm_topic_basic (codes 1xxx and 2xxx): Basic prerequisites for materials modelling, *e.g.*, contents from secondary or undergraduate education.
- Concept otras:mm_topic_computational (codes 3xxx): Computational and numerical aspects of materials modelling.
- Concept otras:mm_topic_data (codes 4xxx): Data science and technology aspects.
- Concept otras:mm_topic_materials (codes 5xxx): Topics related to materials and their properties.
- Concept otras:mm_topic_social (codes 6xxx): Social, economic, and community aspects of materials modelling.
- Concept otras:mm_topic_theoretical (codes 7xxx): Non-computational theoretical aspects of materials modelling.
- Concept otras:mm_topic_interdisciplinary (codes 8xxx): Topics that are best described as belonging to multiple categories at the first hierarchy level.
- Concept otras:mm_topic_side (codes 9xxx): Topics from other disciplines that may be included as relevant side interests in a materials modelling curriculum.

In particular, this taxonomy is used to retrieve training contents, and to indicate relevant areas of interest and fields of knowledge to be used to matchmaking by the *translation router app* of the VIMMP marketplace platform [Hor20a]. OTRAS also permits the specification of topics via CCS, a taxonomy developed by the Association for Computing Machinery [Ass12], and PhySH, developed by the American Physical Society [Ame20].

3.4.3. Operator catalogue for competency specification

An operator specifier, X_2 in Eq. (3.4.1), indicates what sort of activity is enabled by possessing a certain competency. Learning outcomes in course syllabi are typically formulated concisely, *e.g.*, “the students will be able to apply statistical mechanics to problems from fluid phase thermodynamics”. In this example, the operator *is expressed by* the predicate “to apply”. In the interest of the legibility of a syllabus (and the work involved in writing it), generally, a precise definition of the meaning of the operator specifier is not provided – the interpretation is left to the intuition of the reader. Nonetheless, it is in the interest of universities, schools, and other training providers to reach an agreement on a more detailed specification of the semantics associated with a learning outcome formulation; this has aspects of both semantic and pragmatic interoperability, such as where multiple instructors are expected to abide by the same syllabus and/or conduct exams that confirm the success of the learning effort at a specified level.

For this purpose, OTRAS relies on a catalogue of operators disseminated by the German *Kultusministerkonferenz*, facilitating the specification of learning outcomes in the natural sciences in a consistent way [Kul14]. These operators, roughly corresponding to elementary, intermediate, and advanced levels of learning, are complemented in OTRAS by additional operator individuals that are expected to be more adequate for expressing certain competencies that are typically attributed to expert personnel:

- Operator codes 1xx – predominantly used for basic-level competencies: "To name/label" (120), "to outline/present" (130), "to list/give" (140), "to write a lab report/data log" (150), "to sketch" (160), and "to draw" (170).
- Operator codes 2xx – predominantly used for intermediate-level competencies: "To compare" (215), "to deduce" (220), "to estimate" (225), "to analyse and identify" (230), "to apply" (235), "to calculate" (240), "to describe" (245), "to find" (250), "to explain" (255), "to describe and explain" (260), "to formulate" (265), "to derive" (270), "to sort/group/classify" (275), "to test/verify" (280), "to investigate/examine" (285), "to generalize" (290), and "to summarize" (295).
- Operator codes 3xx – predominantly used for advanced-level competencies: "To propose a hypothesis" (320), "to evaluate" (330), "to justify/give reasons" (340), "to comment on/assess" (350), "to prove" (360), "to discuss" (370), "to interpret" (380), and "to plan" (390).
- Operator codes 4xx – predominantly used for expert-level competencies: "To review/evaluate critically" (420), "to advise/manage" (425), "to characterize experimentally" (430), "to document" (435), "to carry out professional work" (440), "to correspond" (445), "to teach" (450), "to plan/project/propose" (455), "to conduct an exam/assessment" (460), "to systematize" (465), "to expand/extend/generalize" (470), "to simplify/reduce" (475), and "to innovate/develop" (480).

Each operator has a three-digit operator code (*e.g.*, 235) and *is expressed by* a concise predicate (*e.g.*, "apply"), while it *is defined by* a more detailed explanation of its meaning; *e.g.*, "use a known idea, equation, principle, theory, or law in a new situation."

3.4.4. Illustration

Selected concepts from OTRAS:

- otras:focus, *i.e.*, a studied object, topic, training objective, or an aspect or constitutive part thereof. Subclasses include otras:learning_outcome, otras:operator_level, otras:specifier, and otras:topic. EMMO-VIPRS alignment:
$$\begin{aligned}\text{otras:focus} &\sqsubseteq \text{evmpo:annotation}\\ &\sqsubseteq \text{emmo-perceptual:Symbolic} \sqcap \exists P.\text{emmo-semiotics:Sign}.\end{aligned}$$
- otras:mm_topic (materials modelling topic), *i.e.*, a topic related to the subject area of materials modelling, understood broadly. EMMO-VIPRS alignment:
$$\begin{aligned}\text{otras:mm\_topic} \sqsubseteq \text{emmo-semiotics:Sign} &\sqcap \exists P.\text{emmo-semiotics:Sign}\\ &\sqcap \text{emmo-perceptual:Symbolic}.\end{aligned}$$
- otras:specifier, *i.e.*, a constitutive element of a learning outcome (competency) description. EMMO-VIPRS alignment: Same as for otras:focus.
- otras:training_service, *i.e.*, a tradeable object (evmpo:tradeable_object) that provides training contents or activities. EMMO-VIPRS alignment:
$$\begin{aligned}\text{otras:training\_service} \sqsubseteq \text{emmo-semiotics:Sign}&\\ &\sqcap \exists S.\text{emmo-holistic:Process}.\end{aligned}$$

- otras:training_unit, *i.e.*, an elementary (part of a) course that is not further subdivided into course parts.
  EMMO alignment: otras:training_unit ⊑ emmo-holistic:Process.

Selected relations (object properties) from OTRAS:

- otras:has_offered_course; points to a course that is offered as part of the activities carried out as a training service. Domain: otras:training_service; range: otras:course.
  EMMO-VIPRS alignment: otras:has_offered_course ⊑ $S \circ P^{-1}$.
- otras:has_specifier; points to an operator, operand, implementation, or example specifier of a learning outcome. Domain: otras:learning_outcome; range: otras:specifier.
  EMMO-VIPRS alignment: otras:has_specifier ⊑ $S^{-1} \circ P^{-1}$.
- otras:is_about; points to an object to which an information content entity refers. This relation is defined to be a subproperty of IAO_0000136, labelled “is about,” from the IAO [Ceu12]. Domain: evmpo:information_content_entity; range: evmpo:marketplace_related_entity.
  EMMO alignment: otras:is_about ⊑ S.
- otras:is_narrower_than; (A is narrower than B): A and B are topics such that if A is a sign for an object, B is also a sign for that object. This relation is defined to be a subproperty of skos:broader [Isa09].[15] Domain: otras:topic; range: otras:topic. Analogous: otras:is_broader_than, which is a subproperty of skos:narrower.
  EMMO-VIPRS alignment: otras:is_narrower_than ⊑ $S \circ S^{-1}$.
- otras:is_part_of_course; points to the course to which the given training unit belongs. Domain: otras:training_unit; range: otras:course.
  EMMO alignment: otras:is_part_of_course ⊑ P.

3.5. VIMMP Communication Ontology (VICO)

The ontology VICO covers metadata on messages exchanged at the virtual marketplace platform and participants that interact at the platform,[16] including end users, model providers, software owners, translators, *etc.* [Hor20a]. Through the LCC ontology, VICO incorporates the ISO 3166 standard for referring to countries and regions [Cel10]. Types of interlocutors (subclasses of vico:interlocutor) are referred to – in accordance with the usual EMMC nomenclature – as consultants, data providers, end users, manufacturers, model providers, software owners, training providers, translators, and guests; the class vico:interlocutor_group contains individuals associated with each of these groups, *e.g.*, vico:software_owner individuals belong to the group vico:IG_SOFTWARE_OWNER.

The communication branch of the class hierarchy is visualized in Fig. 3.10.

[15] In SKOS, the relation is defined the other way around, *i.e.*, “A skos:broader B” means that B is broader. In OTRAS, “A otras:is_narrower_than B” means that A is narrower.

[16] Documented version: VICO v1.3.4, dated 20th November 2020.

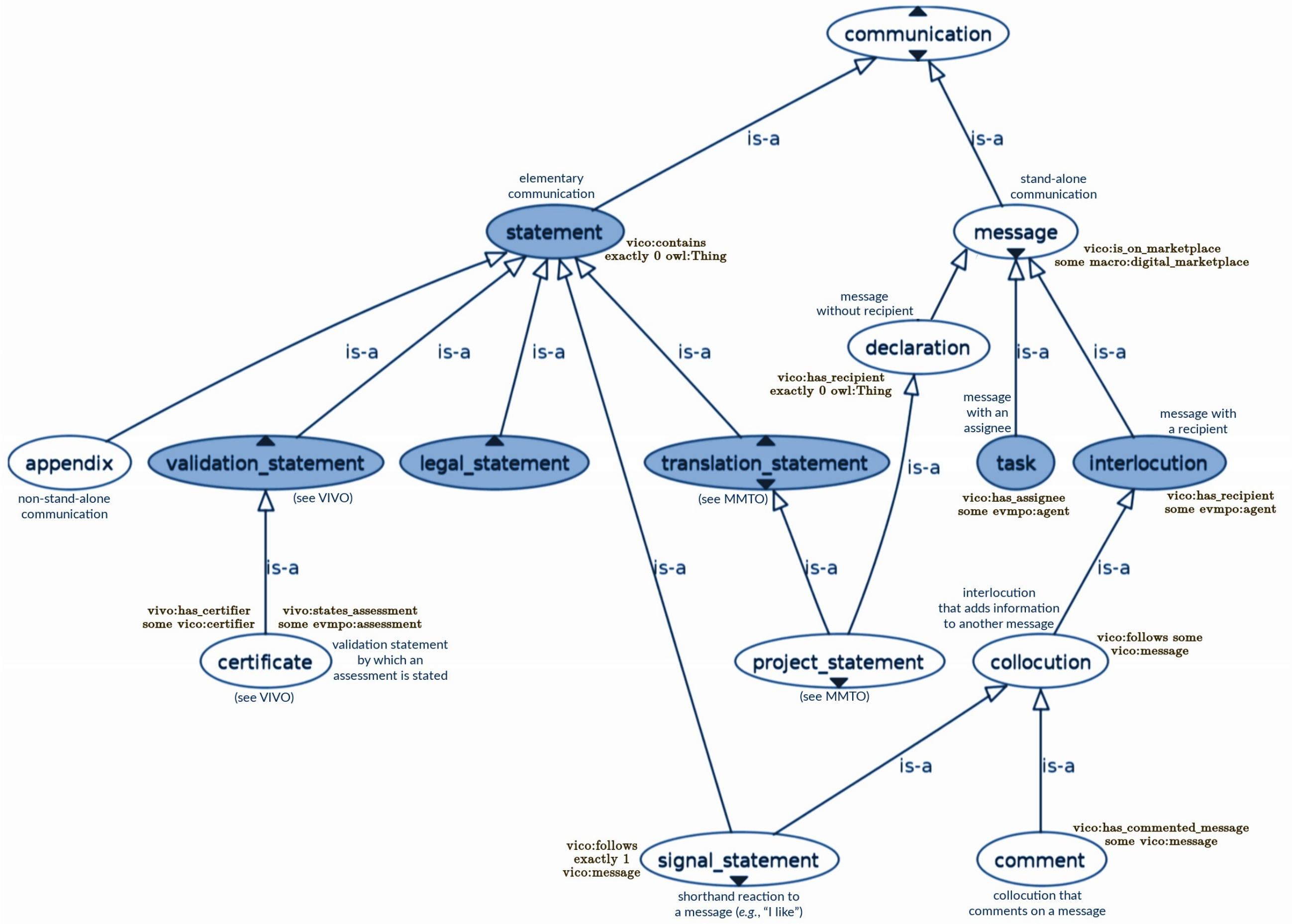


Figure 3.10. Class hierarchy for evmpo:communication and selected subclasses. The diagram was generated using the OWLViz protégé plugin; blue arrows labelled "is-a" denote subsumption ($\sqsubseteq$).

Selected concepts from VICO:

- vico:academic_title, *i.e.*, a titular rank that corresponds to an academic degree.
  EMMO alignment: vico:academic_title $\sqsubseteq$ evmpo:annotation
  $\sqsubseteq$ emmo-perceptual:Symbolic $\sqcap$ $\exists$P.emmo-semiotics:Sign.
- vico:certifier, *i.e.*, an agent who can issue certificates.
  EMMO alignment: vico:certifier $\sqsubseteq$ emmo-semiotics:Interpreter.
- vico:interlocutor_tag, *i.e.*, an interlocutor type that specifies properties of an interlocutor which may co-determine ability/suitability for trading with certain partners at a digital marketplace – indicating the country of residence/registration, whether the described interlocutor is engaged in military or nuclear research, *etc.*
  EMMO alignment: vico:interlocutor_tag $\sqsubseteq$ evmpo:annotation
  $\sqsubseteq$ emmo-perceptual:Symbolic $\sqcap$ $\exists$P.emmo-semiotics:Sign.
- vico:message, *i.e.*, a stand-alone communication (as opposed to an appendix).
  EMMO alignment: vico:message $\sqsubseteq$ emmo-perceptual:Symbolic.
- vico:person, *i.e.*, a stand-alone agent that does not have multiple constituent parts or components each of which could, *e.g.*, act at a digital marketplace by themselves.
  EMMO alignment: vico:person $\sqsubseteq$ emmo-semiotics:Interpreter.

Selected relations (object properties) from VICO:

- vico:contains; (C contains D): D is a proper part of C, where C and D are both communications. Domain: evmpo:communication; range: evmpo:communication. EMMO alignment: vico:contains $\sqsubseteq \mathrm{P}^{-1}$.
- vico:follows; (C follows D): C and D are messages, and C addresses or refers to D. Domain: vico:message; range: vico:message.
  EMMO-VIPRS alignment: vico:follows $\sqsubseteq \mathrm{S} \sqcap$ viprs:is_enabled_by.
- vico:has_affiliation; indicates an institutional affiliation. Domain: vico:person; range: evmpo:institution. EMMO alignment: vico:has_affiliation $\sqsubseteq \mathrm{P}$.
- vico:has_author; points to the agent that has issued the given communication. Domain: evmpo:communication; range: evmpo:agent.
  EMMO-VIPRS alignment: vico:has_author $\sqsubseteq$ viprs:is_enabled_by.
- vico:is_certifier_of; points to a certificate for which a certifier is (co-)responsible, having either issued the certificate or formally approved of its content. Domain: vico:certifier; range: vivo:certificate.
  EMMO-VIPRS alignment: vico:is_certifier_of $\sqsubseteq$ viprs:satisfies_requirement_of.

3.6. VIMMP Software Ontology (VISO)

The aim of VISO[17] is to characterize software tools in the area of materials modelling, especially their features (*i.e.*, capabilities), intended both at the model and solver level, but also their technical requirements, compatibility with other tools and licensing aspects. The concepts defined within this ontology will, first, guide the ingest of information on the VIMMP platform, and, later, allow the users to retrive and compare tools. Below an upper level (viso-general, *cf.* Fig. 3.11) that addresses aspects common to all software, we split VISO into three branches focusing on classes of models: electronic (EL, viso-el), atomistic and mesoscopic (AM, viso-am), and continuum (CO, viso-co) models (*cf.* Figs. 3.12, 3.13 and 3.14). These branches depend on viso-general, but can be loaded independently of the other two siblings. Accordingly, selected major concepts from viso-general are:

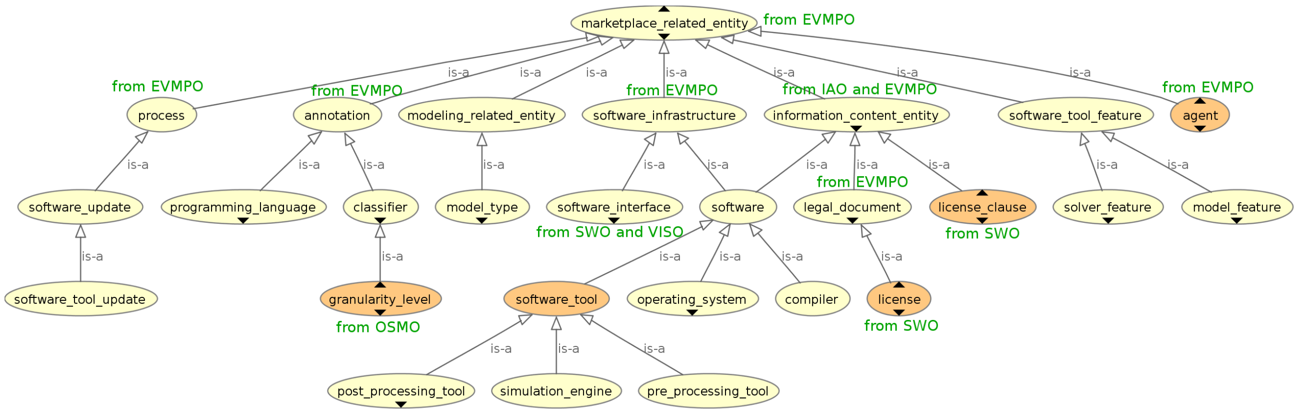


Figure 3.11. Fragment of VISO showing its upper and intermediate classes and their connection to EVMPO and external assets; the diagram was generated using the OWLViz protégé plugin; grey arrows labelled “is-a” denote subsumption ($\sqsubseteq$), *i.e.*, rdfs:subClassOf.

[17] Documented version: VISO v1.0.4, dated 26th November 2020. Significant contributions to VISO development by G. Boccardo, M. Chiricotto, J. D. Elliott are acknowledged [Hor20b].

- viso:software, *i.e.*, a computer program. Its direct (mutually disjoint) subclasses are: viso:software_tool, viso:compiler, viso:operating_system.
  EMMO alignment: viso:software $\sqsubseteq$
  $(\text{emmo-perceptual:Symbolic} \sqcup \text{emmo-manufacturing:Engineered})$.
- viso:programming_language, *i.e.*, a language that can be used to write software.
  EMMO alignment: viso:programming_language $\sqsubseteq$ emmo-perceptual:Language.
- viso:software_tool_feature $\equiv$ (viso:model_feature $\sqcup$ viso:solver_feature), *i.e.*, a capability of a software tool, intended as either a model aspect that can be addressed (viso:model_feature) or as a numerical algorithm which is implemented (viso:solver_feature). Following the approach from RoMM [Deb17], these two classes are disjoint.
- viso:model_type, *i.e.*, a classification of the model, intended as in RoMM [Deb17].
  EMMO alignment: viso:model_type $\sqsubseteq$ emmo-perceptual:Symbolic.
- viso:model_object, *i.e.*, the type of object entering the model and carrying degrees of freedom. Its subclasses in the AM branch (*cf.* Fig. 3.13) include viso-am:interaction_site, viso-am:interaction_surface, viso-am:connected_object; in the EL branch (*cf.* Fig. 3.12) they include viso-el:quantum_object and viso-el:classical_object.
  EMMO alignment: viso:model_object $\sqsubseteq$ emmo-perceptual:Symbolic.
- viso:software_update, *i.e.*, to describe (as text) the changes between versions of a software. In particular, its subclass viso:software_tool_update allows to describe the addition/removal of features from a tool.
  EMMO alignment: viso:software_update $\sqsubseteq$ emmo-holistic:Process.
- viso:software_interface, *i.e.*, an interface between a software and a user or a client (i.e., a program or device). Some sub-classes of this class are taken from the SWO software interface class (swo:SWO_9000050).
  EMMO alignment: viso:software_interface $\sqsubseteq$ emmo-manufacturing:Engineered.
- viso:license, *i.e.*, a regulation of the right to use, modify and distribute something, in this case software. It is declared to be equivalent to the Software Licence class from SWO (swo:SWO_0000002), *cf.* Malone *et al.* [Mal14b].
  EMMO alignment: viso:license $\sqsubseteq$ emmo-perceptual:Symbolic.
- viso:license_clause, *i.e.*, equivalent to the Licence clause class from SWO (swo:SWO_9000005), *cf.* Malone *et al.* [Mal14b].
  EMMO alignment: viso:license_clause $\sqsubseteq$ emmo-perceptual:Symbolic.

Selected relations (object properties) from VISO are:

- viso:has_feature, *i.e.*, points to a (model or solver) feature of a tool.
  Domain: viso:software_tool; range: viso:software_tool_feature.
  EMMO-VIPRS alignment: viso:has_feature $\sqsubseteq$ viprs:enables_entity_containing.
- viso:is_compatible_with, *i.e.*, is able to exchange information directly, with no need to interface. Domain and range: viso:software_tool.
  EMMO-VIPRS alignment: viso:is_compatible_with $\sqsubseteq$ viprs:can_cooccur_with.

- viso:involves, *i.e.*, (X involves Y) means that there is a mathematical expression or an algorithmic formulation of X that contains Y.
  Domain: viso:software_tool_feature $\sqcup$ viso:modeling_related_entity; range: vov:variable $\sqcup$ vov:function $\sqcup$ viso:model_object.
  EMMO alignment: viso:involves $\sqsubseteq$ P^{-1}.
- viso:is_tool_for_model, *i.e.*, associates tools with models. Domain: viso:software_tool; range: viso:model_type. EMMO-VIPRS alignment: viso:is_tool_for_model $\sqsubseteq$ viprs:enables_entity_containing.
- viso:requires, *i.e.*, relates a tool to libraries and/or operating systems.
  Domain: viso:software_tool; range: viso:software.
  EMMO-VIPRS alignment: viso:requires $\sqsubseteq$ viprs:n_lnoc_or_roc.
- viso:is_modelling_twin_of, *i.e.*, relates two objects that (despite being possibly distinct individuals) are equivalent from the modelling point of view.
  EMMO-VIPRS alignment: viso:is_modelling_twin_of $\sqsubseteq$ viprs:immanent_relation.

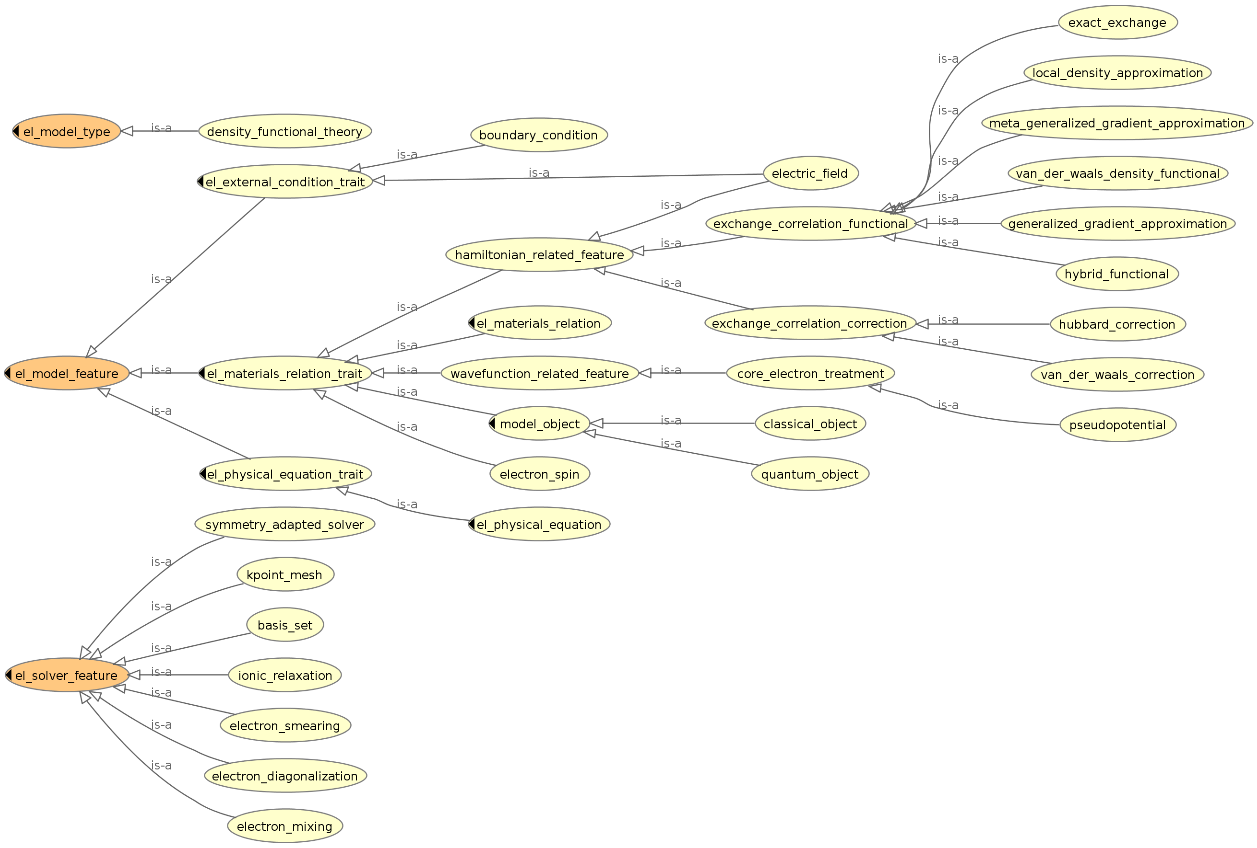


Figure 3.12. Branch of VISO for electronic models (viso-el); the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption ($\sqsubseteq$), *i.e.*, rdfs:subClassOf.

Below viso:general, the EL, AM and CO branches of VISO expand on the categories viso:model_feature, viso:solver_feature and viso:model_type (*cf.* Figs. 3.12, 3.13 and 3.14). The three branches have a common structure, in that the subclasses of viso:model_feature are classified into (non-disjoint) classes viso:materials_relation_trait, viso:physical_equation_trait, and viso:external_condition_trait. For clarity, we systematically use "trait" here, and not "aspect", since the latter keyword has a different and well defined role within OSMO and MODA.

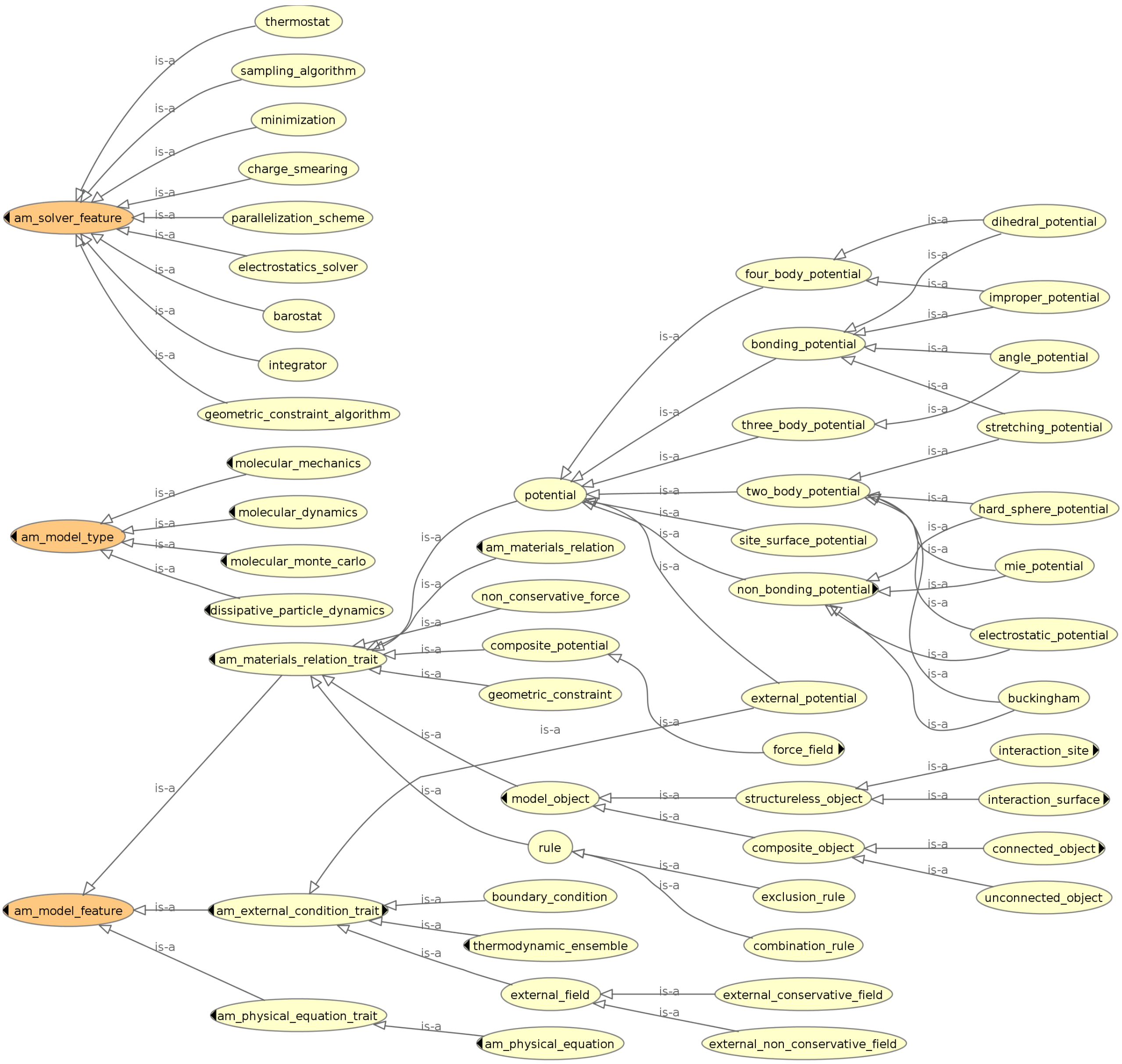


Figure 3.13. Branch of VISO for atomistic-mesoscopic models (viso-am); the diagram was generated using the OWLViz protégé plugin; grey arrows labelled “is-a” denote subsumption (⊑), *i.e.*, rdfs:subClassOf.

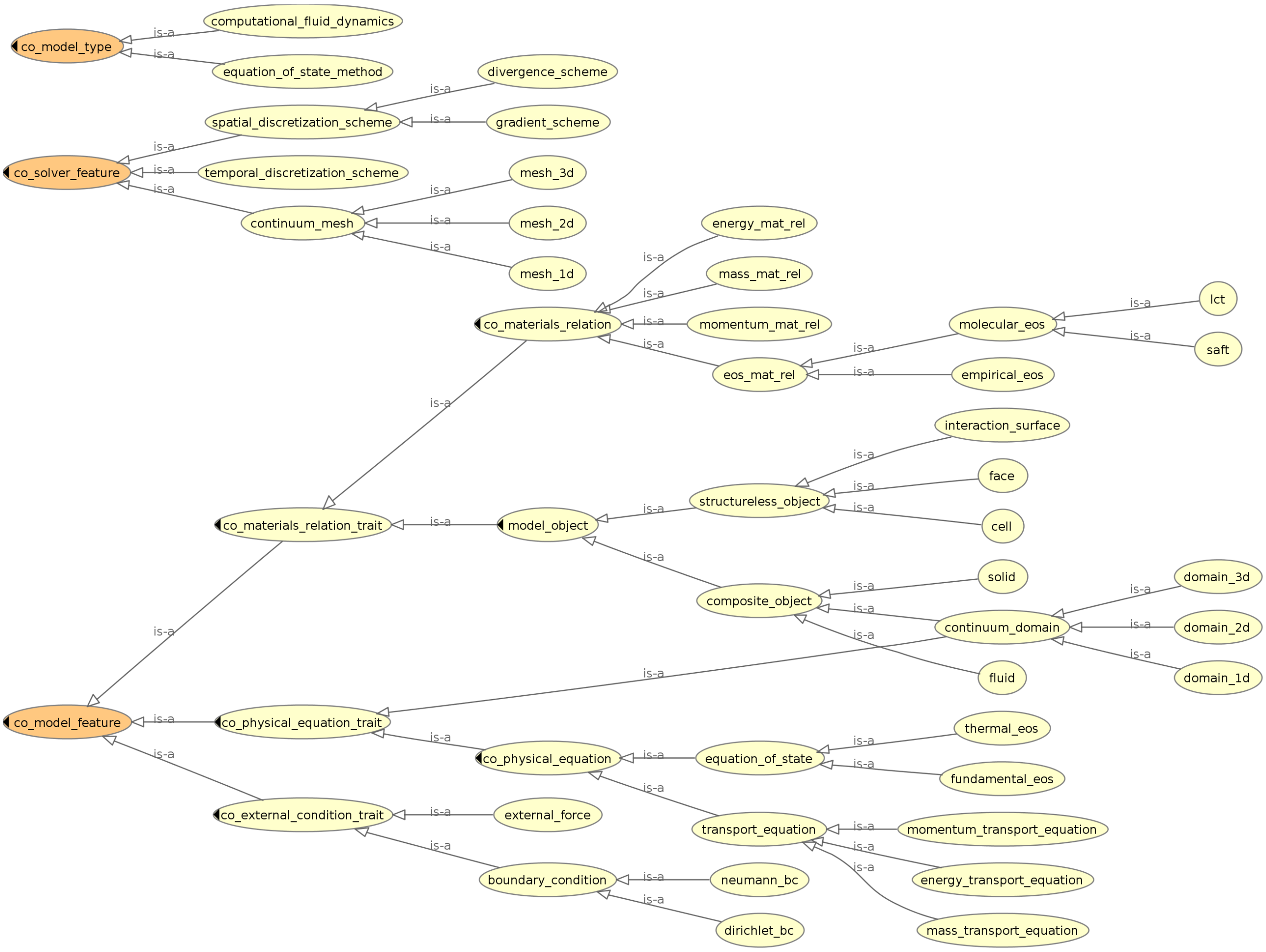


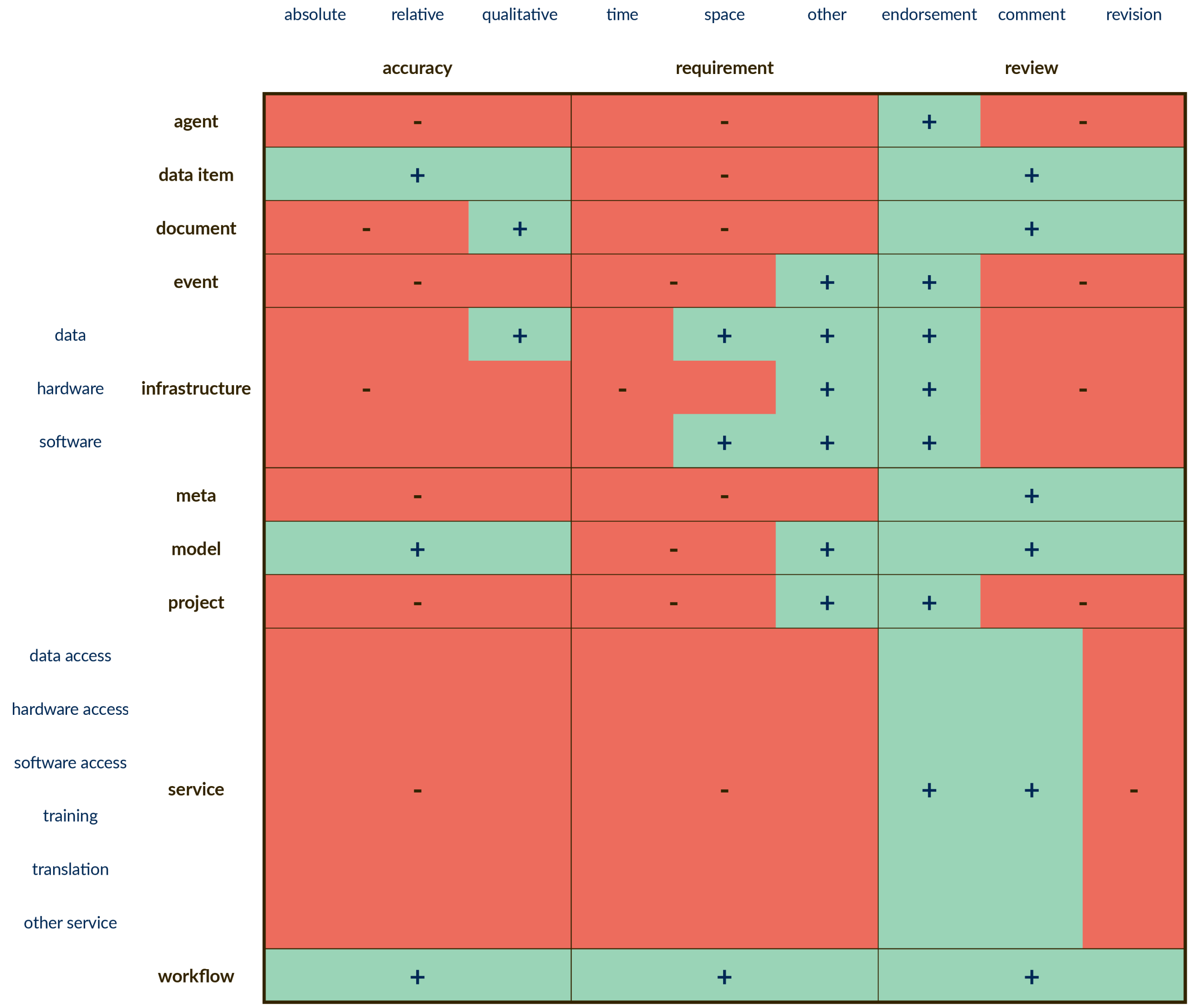
Figure 3.14. Branch of VISO for continuum models (viso-co); the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption ($\sqsubseteq$), *i.e.*, rdfs:subClassOf.

3.7. VIMMP Validation Ontology (VIVO)

The ontology VIVO categorizes assessments (*i.e.*, evaluations) of computational resource requirements and benchmarking as well as customer feedback on various kinds of marketplace-relevant entities,[18] which can be provided subsequent to transactions at the VIMMP marketplace [Hor20a]. Thereby, users support each other mutually, evaluating contents and providers, while the marketplace platform itself remains neutral and equally open and accessible to everybody. A matrix with subclasses of evmpo:assessment, which indicates how marketplace users can evaluate what sort of objects, is shown in Fig. 3.15. Rows correspond to classes of entities that are subjected to an assessment, such that, *e.g.*, a vivo:data_infrastructure_assessment is an evmpo:assessment that vivo:evaluates an evmpo:data_infrastructure, and a vivo:meta_assessment is an evmpo:assessment that vivo:evaluates an evmpo:assessment. Columns corresponds to different ways in which entities can be evaluated, *e.g.*, by providing feedback on the observed relative quantitative accuracy (vivo:relative_accuracy_assessment) or by issuing a recommendation to other users (evmpo:endorsement_assessment). Not all theoretically conceivable combinations are allowed – *e.g.*, memory requirements can be stated for software, but not for projects.

[18] Documented version: VIVO v1.1.6, dated 20th November 2020.

Using VIVO, in particular, error analyses and estimates can be attributed to models, simulation workflows, and to data items obtained from repositories or other platforms.



		absolute	relative	qualitative	time	space	other	endorsement	comment	revision
		accuracy			requirement			review		
	agent	-	-	-	-	-	-	+	-	-
	data item	+	+	+	-	-	-	+	+	+
	document	-	-	+	-	-	-	+	+	+
	event	-	-	-	-	-	+	+	-	-
data	infrastructure	-	-	+	-	+	+	+	-	-
hardware		-	-	-	-	-	+	+	-	-
software		-	-	-	-	+	+	+	-	-
	meta	-	-	-	-	-	-	+	+	+
	model	+	+	+	-	-	+	+	+	+
	project	-	-	-	-	-	+	+	-	-
data access	service	-	-	-	-	-	-	+	+	-
hardware access		-	-	-	-	-	-	+	+	-
software access		-	-	-	-	-	-	+	+	-
training		-	-	-	-	-	-	+	+	-
translation		-	-	-	-	-	-	+	+	-
other service		-	-	-	-	-	-	+	+	-
	workflow	+	+	+	+	+	+	+	+	+

Figure 3.15. Matrix of permitted (+) and prohibited (–) types of assessments.

Selected concept definitions from VIVO:

- vivo:assertion, *i.e.*, a claim or proposition (*e.g.*, as part of an assessment). Subclasses include vivo:accuracy_assertion, evmpo:material_property_information, and vivo:requirement_assertion, *cf.* Fig. 3.15.
EMMO alignment: vivo:assertion $\sqsubseteq$ emmo-perceptual:Symbolic.
- vivo:certificate, *i.e.*, a validation statement by which an assessment is stated.
EMMO alignment: vivo:certificate $\sqsubseteq$ emmo-perceptual:Symbolic.
- vivo:computational_time_requirement, *i.e.*, a requirement assessment by which the computational (CPU time) requirements of a simulation workflow are evaluated.
EMMO-VIPRS alignment:

$$\begin{aligned}&\text{vivo:computational\_time\_requirement}\\ &\quad\sqsubseteq \text{emmo-semiotics:Sign} \sqcap \exists(\mathrm{S}\circ\mathrm{S}).\text{emmo-semiotics:Semiosis}\\ &\qquad\sqcap \exists\mathrm{S}.(\text{emmo-semiotics:Sign} \sqcap \text{emmo-perceptual:Symbolic}).\end{aligned}$$

- vivo:material_property_information, *i.e.*, an assertion that refers to a material property by means of the relation vivo:refers_to_mp, see below.
  EMMO-VIPRS alignment:
  $$\begin{aligned} &\text{vivo:material\_property\_information} \\ &\quad \sqsubseteq \text{emmo-semiotics:Sign} \sqcap \exists\text{S.emmo-properties:Property} \\ &\qquad \sqcap \exists\text{S.emmo-physicalistic:Material} \\ &\qquad \sqcap \exists(\text{S}\circ\text{S}).\text{emmo-physicalistic:Material}. \end{aligned}$$
- vivo:relative_deviation, *i.e.*, an accuracy assertion in which the relative magnitude of an error or uncertainty is given, normalized by the absolute magnitude of the value to which the assertion refers. Analogous: vivo:absolute_deviation.
  EMMO alignment: vivo:relative_deviation $\sqsubseteq$ emmo-math:Mathematical.
- vivo:revision, *i.e.*, a review assessment by which a replacement for the entity under review is recommended.
  EMMO alignment: vivo:revision $\sqsubseteq$ evmpo:assessment $\sqsubseteq$ emmo-semiotics:Sign.
- vivo:unit, *i.e.*, a unit that can be expressed as a scalar multiple of an algebraic combination of SI units. This concept is the same as qudt:Unit and emmo-metrology:ReferenceUnit [Fra20, Zha17].
  EMMO alignment: vivo:unit $\equiv$ emmo-metrology:ReferenceUnit.

Selected relations (object properties) from VIVO:

- vivo:evaluates; points to the object evaluated by an assessment. Domain: evmpo:assessment; range: evmpo:marketplace_related_entity.
  EMMO alignment: vivo:evaluates $\sqsubseteq$ S.
- vivo:has_assertion; points to an assertion made within an assessment. Domain: evmpo:assessment; range: vivo:assertion.
  EMMO alignment: vivo:has_assertion $\sqsubseteq$ P^{-1}.
- vivo:has_error_statement; points to an accuracy assertion contained within a material property information. Domain: vivo:material_property_information; range: vivo:accuracy_assertion.
  EMMO alignment: vivo:has_error_statement $\sqsubseteq$ P^{-1}.
- vivo:has_unit; points to the unit in which any numerical contents of an assertion are given. Domain: vivo:assertion; range: vivo:unit.
  EMMO-VIPRS alignment: vivo:has_unit $\sqsubseteq$ viprs:has_symbolic_part_of_sign.
- vivo:is_quantity_kind; points to the physical property characterization following QUDT [Zha17], *e.g.*, qudt:MassUnit or qudt:EnergyAndWorkUnit. Domain: vivo:assertion; range: qudt:QuantityKind.
  EMMO alignment: vivo:is_quantity_kind $\sqsubseteq$ S^{-1}.
- vivo:refers_to_mp; points to the material property (MP) to which a material assertion refers. Domain: vivo:material_assertion; range: evmpo:material_property.
  Analogous: vivo:refers_to_material.
  EMMO alignment: vivo:refers_to_mp $\sqsubseteq$ vivo:refers_to_property $\sqsubseteq$ S.
- vivo:states_assessment; points to an assessment contained within a certificate. Domain: vivo:certificate; range: evmpo:assessment.

EMMO-VIPRS alignment: vivo:states_assessment $\sqsubseteq \mathrm{P}^{-1} \sqcap$ viprs:enables.

3.8. VIMMP Ontology of Variables (VOV)

The purpose of VOV[19] is to organize the variables (in a broad sense, including constants) that appear in modelling and simulation, and to connect them to models and algorithms in which they are involved as well as to model objects (*e.g.*, entities entering a simulation, such as sites, rigid bodies) to which they are attached. VOV can be used in connection with VISO and OSMO to further specify models, algorithms and workflows.

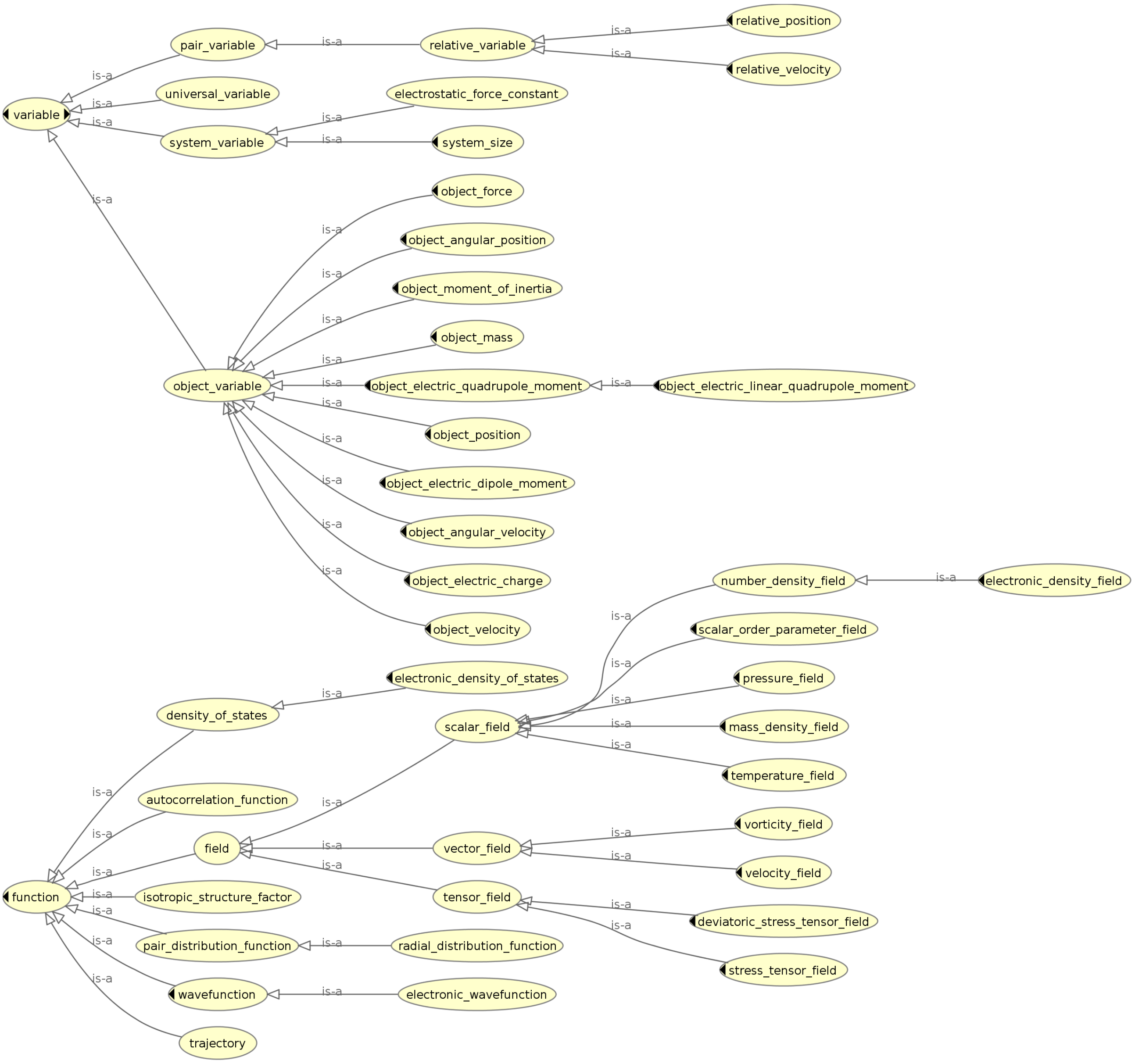


Figure 3.16. Fragment of VOV showing selected subclasses of vov:variable and the subclasses of vov:function; the diagram was generated using the OWLViz protégé plugin; grey arrows labelled "is-a" denote subsumption ($\sqsubseteq$), *i.e.*, rdfs:subClassOf.

The main concepts from VOV are:

[19] Documented version: VOV v1.0.5, dated 26th November 2020. Significant contributions to VOV development by J. D. Elliott are acknowledged.

- $\mathrm{vov{:}variable}$, *i.e.*, a variable in the mathematical sense – a symbol that stands for a quantity in a mathematical expression.
  EMMO alignment: $\mathrm{vov{:}variable} \sqsubseteq \mathrm{emmo\text{-}math{:}Variable}$.
- $\mathrm{vov{:}function}$, *i.e.*, a relation between two or more variables (*e.g.*, the radial distribution function, the energy density of states); it can be defined via a mathematical equation or via tabulated values. Its subclasses (*cf.* Fig. 3.16) include $\mathrm{vov{:}field}$.
  EMMO alignment: $\mathrm{vov{:}function} \sqsubseteq \mathrm{emmo\text{-}math{:}Graphical}$.

Variables in VOV can be classified according to three main criteria: By their scope (vov:object_variable, vov:pair_variable, vov:system_variable, vov:universal_variable), their rank (vov:scalar_variable, vov:vector_variable, vov:tensor_variable) or the kind of quantity (vov:mass, vov:energy, ...), for which qudt:QuantityKind is used [Zha17]. In Fig. 3.15, we show the splitting of vov:variable according to scope and the subclasses of vov:function. Selected relations (object properties) from VOV are:

- vov:has_attached_variable, *i.e.*, points to a variable that is carried by/attached to an object. Its subproperties include vov:has_mass vov:has_coordinates and vov:has_velocity.
  Domain: viso:model_object; range:vov:object_variable.
- vov:has_attached_function, *i.e.*, points to a function that is carried by/attached to an object. Its subproperties include vov:has_velocity_field and vov:has_wavefunction.
  Domain: viso:model_object; range:vov:function.

Note that both relations specify viso:involves, *i.e.*: vov:has_attached_variable $\sqsubseteq$ viso:involves and vov:has_attached_function $\sqsubseteq$ viso:involves.

4. List of referenced external semantic assets

- **CCS: Computing Classification System** [Ass12]; a taxonomy of topics maintained by the Association for Computing Machinery (ACM). It can be used to specify topics, in addition to those included in the materials modelling topic taxonomy from OTRAS.
- **CCSO: Course, Curriculum, and Syllabus Ontology** [Kat18]. This ontology provides key concepts and relations used in OTRAS to describe training events and topics, including ccso:Course, ccso:LearningOutcome, ccso:Syllabus, and ccso:Topic.
- **DCMES: Dublin Core Metadata Element Set** [Dub20]. The DCMES concept dcterms:W3C-DTF is used for communication-related timestamps in VICO.
- **EDAM: EMBRACE Data and Methods** [Iso13]; file format ontology used by MACRO.
- **EMMO: European Materials and Modelling Ontology** [Fra20]; top-level ontology to which the concepts and relations from the VIMMP domain ontologies are aligned.
- **ETCT: EMMC Translation Case Template** [Emm17]. The ETCT contains forms by which a translation case in materials modelling can be described in a structured way; the MMTO is the ontology version of the ETCT.

- **ETG: EMMC Translators' Guide** [Hri19]. This document outlines recommendations and agreements on good practice, contributing to *pragmatic interoperability* in materials modelling translation, which is covered here by the MMTO.
- **FIBO: Financial Industry Business Ontology** [Ben14], developed by the Enterprise Data Management Council. The Currency Amount Ontology (CAO), a module from FIBO, is used by the MMTO for the object content of the business-case aspect mmto:bca_currency, indicating the currency used for budgeting.
- **IAO: Information Artifact Ontology** [Ceu12]. One of the fundamental categories from the EVMPO, "information content entity," is taken from the IAO. Accordingly, the IAO is frequently referenced, *e.g.*, for training-related documents in OTRAS.
- **iCal: W3C iCalendar Ontology** [Con05]. This ontology is widely used for calendar applications; here, it is applied to the "calendar event" fundamental category.
- **KMK Operator Catalogue** [Kul14]: English-language learning-outcome operator designations in natural sciences, provided by the German *Kultusministerkonferenz* (KMK); OTRAS relies on this catalogue for competency specification.
- **LCC: Languages, Countries, and Codes.** This is an implementation of the ISO 3166 standard [Cel10]. In VICO, the LCC country representations module is used, stating $\mathrm{vico{:}country} \equiv \mathrm{lcc\text{-}cr{:}Country}$.
- **MODA: Model Data** [Cen18]. Semi-formalized metadata standard for simulation workflows that serves as the basis for OSMO, the ontology version of MODA.
- **QUDT: Quantities, Units, and Datatypes** [Zha17]. VOV uses the relation $\mathrm{qudt{:}hasQuantityKind}$, connecting variables to a physical quantity, as well as individual physical quantities defined by QUDT and its quantity-kind module, to characterize variables. QUDT units and quantity kinds are also referred to in VIVO.
- **PaaSPort Ontology** [Bas18]: PaaS stands for "platform as a service." This ontology was developed for semantic-interoperability purposes by the PaaSPort Marketplace project, which was funded from the 7th EU Framework Programme and bears certain similarities to VIMMP. The PaaSPort ontology can be used to complement MACRO.
- **PhySH: Physics Subject Headings** [Ame20], a taxonomy developed by the American Physical Society; like CCS, it can be used for topics that are absent from OTRAS.
- **RoMM: Review of Materials Modelling** [Deb17]. This compendium of the field of materials modelling is the basic point of departure for a community-governed line of work in metadata standardization to which the present ontologies also belong. RoMM is a book for human readers, not a machine-processable semantic asset, which precludes an immediate integration of RoMM concepts into ontologies through semantic web references, *i.e.*, by pointing to IRIs. However, particularly through OSMO, the VIMMP ontologies follow the approach from RoMM closely.
- **SKOS: Simple Knowledge Organization System** [Isa09]. The SKOS meta-ontology deals with concept schemes and relations between concepts. In VIMMP, it is used to specify *codelists* (*i.e.*, lists of objects) for classes that contain a finite, well-defined set of individuals; *e.g.*, this is applied to the taxonomy of topics from OTRAS.
- **SWO: Software Ontology** [Mal14b]. VISO uses the SWO to describe software interfaces; for licensing, VISO connects to the license module of the SWO.

References

[All11] D. Allemang and J. Hendler, *Semantic Web for the Working Ontologist*, 2nd edn., Morgan Kaufmann (Waltham), **2011**.

[Ame20] American Physical Society, *Physics Subject Headings*, https://physh.aps.org/, APS, **2020**.

[Ass12] Association for Computing Machinery, *The 2012 ACM Computing Classification System*, https://www.acm.org/publications/class-2012/, ACM, **2012**.

[Baa17] F. Baader, I. Horrocks, C. Lutz, and U. Sattler, *An Introduction to Description Logic*, Cambridge University Press, **2017**.

[Bar12] L. G. Barrientos, E. R. Cruz Sosa, and P. E. García Castro, “Considerations of e-commerce within a globalizing context,” *International Journal of Management and Information Systems* 16(1), 101–110, **2012**.

[Bas18] N. Bassiliades, M. Symeonidis, P. Gouvas, E. Kontopoulos, G. Meditskos, and I. Vlhavas, “PaaSPort semantic model: An ontology for a platform-as-a-service semantically interoperable marketplace,” *Data & Knowledge Engineering* 113, 81–115, **2018**.

[Ben14] M. Bennett, “Reuse of semantics in business applications,” in *Ontologies in Conceptual Modeling and Information Systems Engineering*, CEUR-WS (Aachen), **2014**.

[Ber15] F. Berto and N. Plebani, *Ontology and Metaontology: A Contemporary Guide*, Bloomsbury Academic (London), **2015**.

[Cel10] J. Celko, *Data, Measurements and Standards in SQL*, Morgan Kaufmann (Burlington, MA, USA), **2010**.

[Cen18] CEN-CENELEC Management Centre, *Materials Modelling: Terminology, Classification and Metadata*, CEN workshop agreement 17284, CEN (Brussels), **2018**.

[Ceu12] W. Ceusters, “An information artifact ontology perspective on data collections and associated representational artifacts,” *Studies in Health Technology and Informatics* 180, 68–72, **2012**.

[Con05] D. Connolly, L. Miller, *RDF Calendar: An Application of the Resource Description Framework to iCalendar Data*, W3C interest group note, https://www.w3.org/TR/rdfcal/, W3C, **2005**.

[Deb17] A. F. de Baas (ed.), *What Makes a Material Function? Let me Compute the Ways*, EU Publications Office (Luxembourg), **2017**.

[Dep18] Department for Business, Energy & Industrial Strategy, *Industrial Strategy: Building a Britain Fit for the Future*, white paper, UK Government, **2018**.

[Dub20] Dublin Core Metadata Initiative, *DCMI Metadata Terms*, https://www.dublincore.org/specifications/dublin-core/dcmi-terms/, **2020**.

[Eco75] U. Eco, *Trattato di Semiotica Generale*, Bompiani (Milano), **1975**.

[Emm17] EMMC Coordination and Support Action, *EMMC Translation Case Template*, https://emmc.info/emmc-translation-case-template/, **2017**.

[Emm18] EMMC Coordination and Support Action, *Report on Business Related Quality Attributes for Industry Integration of Materials Modelling*, EMMC-CSA project deliverable 6.3, **2018**.

[Fra20] J. Francisco Morgado, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, J. Friis, and A. de Baas, “Mechanical testing ontology for digital-twins: A roadmap based on EMMO,” p. 3 in *SeDiT 2020: Semantic Digital Twins 2020*; CEUR-WS (Aachen), **2020**; *cf.* E. Ghedini *et al.*, “EMMO: An ontology for applied science,” https://emmc.info/emmo-info/, **2020**.

[Ghe20] E. Ghedini, personal communication, **2020**.

[Gyg20] G. Gygli and J. Pleiss, “Simulation foundry: Automated and FAIR molecular modeling,” *Journal of Chemical Information and Modeling* 60(4), 1922–1927, **2020**.

[Hor20a] M. T. Horsch, S. Chiacchiera, M. A. Seaton, I. T. Todorov, K. Šindelka, M. Lísal, B. Andreon, E. Bayro Kaiser, G. Mogni, G. Goldbeck, R. Kunze, G. Summer, A. Fiseni, H. Brüning, P. Schiffels, and W. L. Cavalcanti, “Ontologies for the Virtual Materials Marketplace,” *KI – Künstliche Intelligenz* 34(3), 423–428, **2020**.

[Hor20b] M. T. Horsch, C. Niethammer, G. Boccardo, P. Carbone, S. Chiacchiera, M. Chiricotto, J. D. Elliott, V. Lobaskin, P. Neumann, P. Schiffels, M. A. Seaton, I. T. Todorov, J. Vrabec, and W. L. Cavalcanti, "Semantic interoperability and characterization of data provenance in computational molecular engineering," *Journal of Chemical & Engineering Data* 65(3), 1313–1329, **2020**.

[Hor21] M. T. Horsch, S. Chiacchiera, W. L. Cavalcanti, B. Schembera, *Data Technology in Materials Modelling*, Springer (Cham), **2021**.

[Hri19] D. Hristova-Bogaerds, P. Asinari, N. Konchakova, L. Bergamasco, A. Marcos Ramos, G. Goldbeck, D. Höche, O. Swang, and G. J. Schmitz, *EMMC Translators' Guide*, doi:10.5281/zenodo.3552260, **2019**.

[Hut04] M. Huth and M. Ryan, *Logic in Computer Science: Modelling and Reasoning about Systems*, 2nd edn., Cambridge University Press, **2004**.

[Isa09] A. Isaac and E. Summers, *SKOS Simple Knowledge Organization System Primer*, W3C working group note, W3C, **2009**.

[Iso13] J. Ison, M. Kalaš, I. Jonassen, D. Bolser, M. Uludağ, H. McWilliam, J. Malone, R. López, S. Pettifer, and P. Rice, "EDAM: An ontology of bioinformatics operations, types of data and identifiers, topics and formats," *Bioinformatics* 29(10), 1325–1332, **2013**.

[Kat18] E. Katis, H. Kondylakis, G. Agathangelos, and V. Kostas, "Developing an ontology for curriculum & syllabus," pp. 55–59 in *The Semantic Web: ESWC 2018 Satellite Events*, Springer (Cham), **2018**.

[Kul14] Kultusministerkonferenz, *Operatoren für die naturwissenschaftlichen Fächer in englischer Sprache an den Deutschen Schulen im Ausland*, version 265, KMK (Bonn), **2014**.

[Mac18] J. Machač, F. Steiner, and J. Tupa, "Product life cycle risk management," pp. 51–72 in C. F. Oudoza (ed.), *Risk Management Treatise for Engineering Professionals*, IntechOpen, **2018**.

[Mai17] J. Maier, *Made Smarter. Review 2017*, technical report, Department for Business, Energy & Industrial Strategy, UK Government, **2017**.

[Mal14a] N. Malankowski and J. C. Brandt, *Innovations- und Effizienzsprünge in der chemischen Industrie*, report S-2014-685-1, VDI-Technologiezentrum, **2014**.

[Mal14b] J. Malone, A. Brown, A. L. Lister, J. Ison, D. Hull, H. Parkinson, and Robert Stevens, "The Software Ontology (SWO): A resource for reproducibility in biomedical data analysis, curation and digital preservation," *Journal of Biomedical Semantics* 5, 25, **2014**.

[Pei91] C. S. Peirce, *Peirce on Signs: Writings on Semiotic*, University of North Carolina Press (Chapel Hill), **1991**.

[Pri13] K. H. Pries and J. M. Quigley, *Reducing Process Costs with Lean, Six Sigma, and Value Engineering Techniques*, CRC (Boca Raton), **2013**.

[Ru15] G. Rutkai and J. Vrabec, "Empirical fundamental equation of state for phosgene based on molecular simulation data," *Journal of Chemical & Engineering Data* 60(10), 2895–2905, **2015**.

[Ste19] S. Stephan, M. Thol, J. Vrabec, and H. Hasse, "Thermophysical properties of the Lennard-Jones fluid: Database and data assessment," *Journal of Chemical Information and Modeling* 59(10), 4248–4265, **2019**.

[Var96] A. C. Varzi, "Parts, wholes, and part-whole relations: The prospects of mereotopology," *Data & Knowledge Engineering* 20(3), 259–286, **1996**.

[Zha17] X. Zhang, K. Li, C. Zhao, and Dongyu Pan, "A survey on units ontologies: Architecture, comparison and reuse," *Program: Electronic Library and Information Systems* 51(2), 193–213, **2017**.