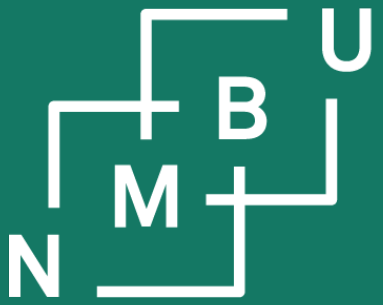


Norges miljø- og biovitenskapelige universitet

Documentation of epistemic metadata by a mid-level ontology of cognitive processes

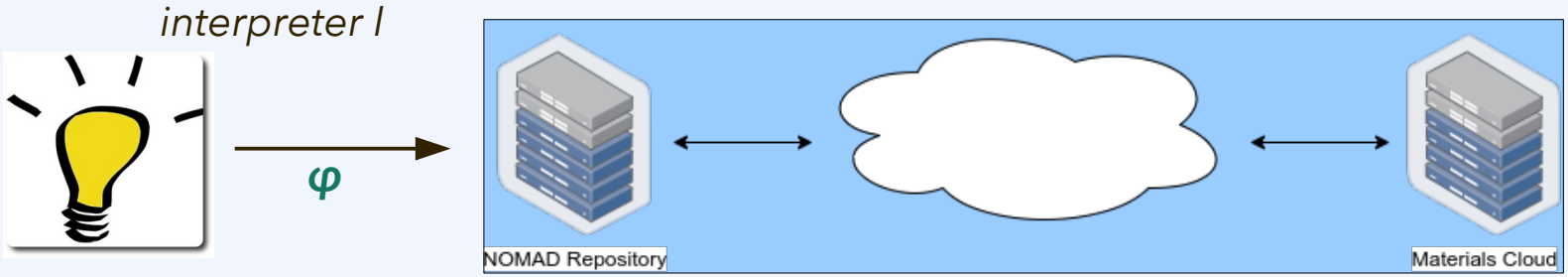
Martin Thomas Horsch[1] and Björn Schembera[2]

[1]Norwegian University of Life Sciences, Department of Data Science
[2]University of Stuttgart, Institute of Applied Analysis and Numerical Simulation

CAOS VI @ JOWO 2022 18th August 2022 Jönköping

Communication of knowledge

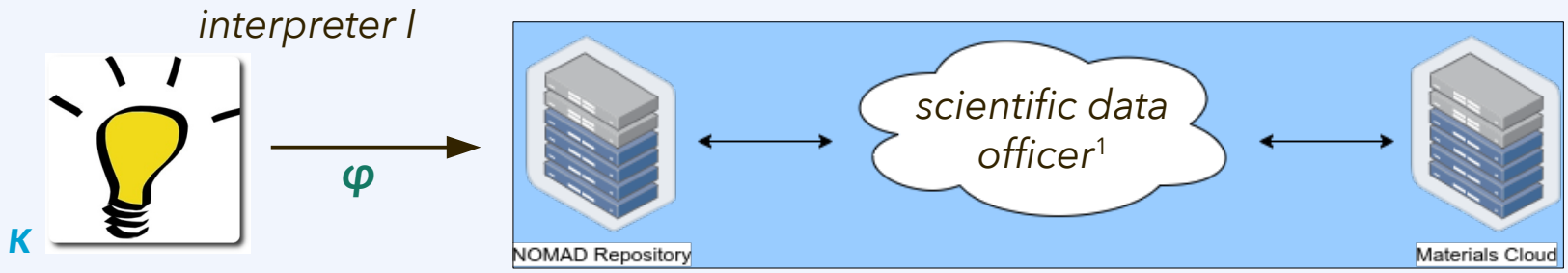
Research data infrastructures store and exchange scientific knowledge.



- An interpreter has made **knowledge claim φ** on the basis of dataset δ.
- **φ** is a **justifiably tenable** proposition, judging by its **epistemic grounding**.

Communication of knowledge

Research data infrastructures store and exchange scientific knowledge.



- An interpreter has made **knowledge claim φ** on the basis of dataset δ.
- The **provenance** of claim φ and dataset δ is that they come **from process κ**.
- φ is a **justifiably tenable** proposition, judging by its **epistemic grounding**.
- "We," *e.g.*, a scientific data officer[1] of the research data infrastructure, have a **justified true belief** in the accuracy of the **provenance documentation κ**.

[1]B. Schembera, J. M. Durán, *Philos. Technol.* **33**: 93–115, doi:10.1007/s13347-019-00346-x, **2019**.

European digitalization platforms

Epistemic opacity (Humphreys, 2011): A cognitive "process is **epistemically opaque** relative to a cognitive agent *X* at time *t* just in case *X* does not know at *t* all of the **epistemically relevant elements** of the process."

European AI Act proposal: "To address the **opacity** that may make certain AI systems **incomprehensible to or too complex for natural persons**, a certain degree of transparency should be required for high-risk AI systems.[1] […] High-risk AI systems should therefore be accompanied by **relevant documentation**".

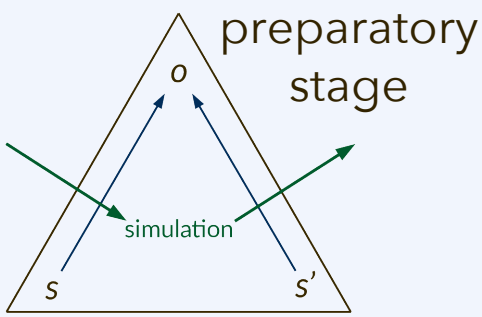
[1]Systems with "high risk" include "safety components" related to "water, gas, heating, and electricity."

Epistemic metadata:

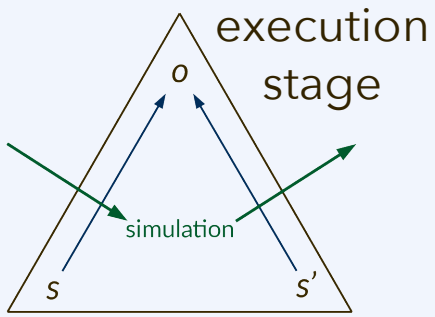
a) "what **knowledge claim (KC) *φ*** has been formulated?,"
b) "where do the data and the claim come from?" (**provenance**),
c) "what **validity claim (VC)** was made about *φ*?,"
d) "why should we accept any of this?" (**grounding**).

Documentation of cognitive processes

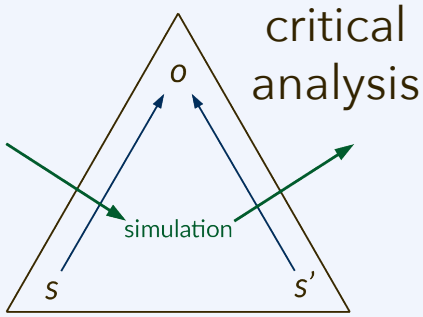
Epistemic opacity (Humphreys, 2011): A cognitive "process is **epistemically opaque** relative to a cognitive agent *X* at time *t* just in case *X* does not know at *t* all of the **epistemically relevant elements** of the process."



ex ante
Planning



in actu
Steering



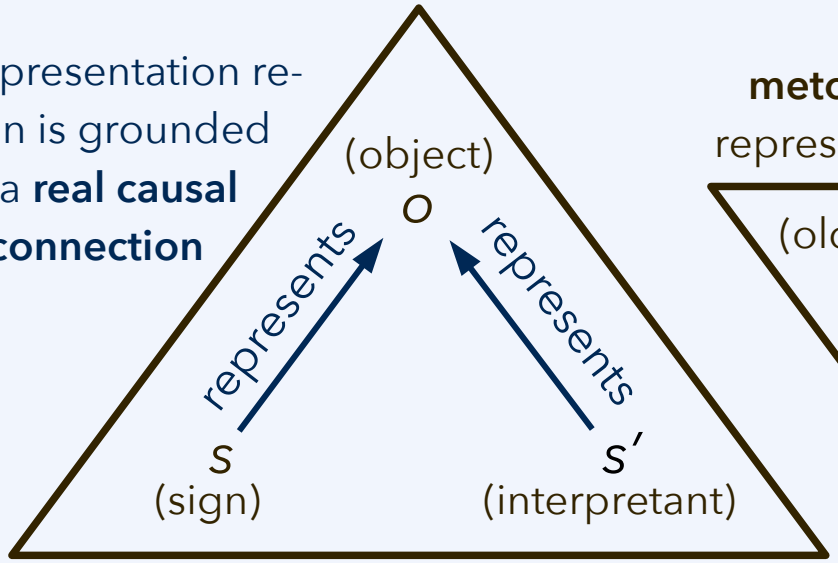
ex post
Evaluation

a) "what **knowledge claim (KC) *φ*** has been formulated?,"
b) "where do the data and the claim come from?" (**provenance**),
c) "what **validity claim (VC)** was made about *φ*?,"
d) "why should we accept any of this?" (**grounding**).

Peircean semiotics and mereosemiotics

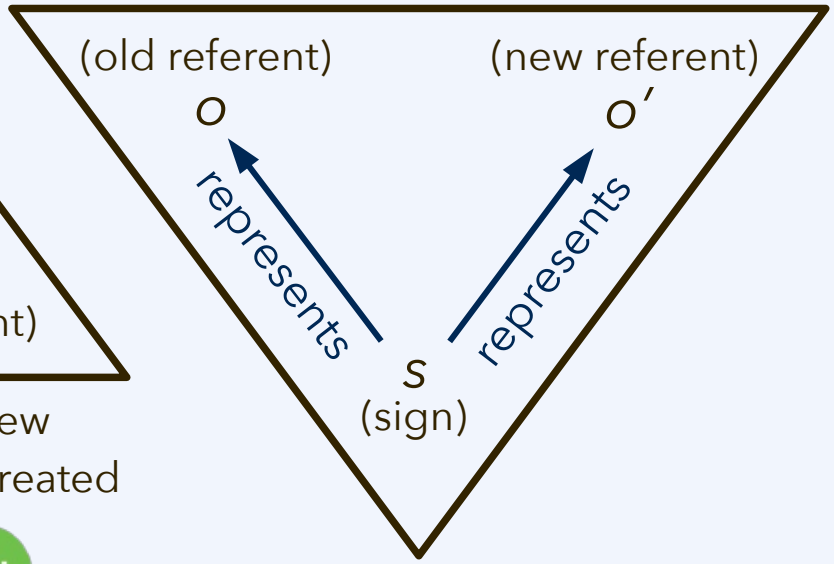
Peircean semiotics: By using a sign (1st) for an object (2nd), a "Third" is created.

the representation relation is grounded in a **real causal connection**



semiosis, a process by which a new representamen, the interpretant, is created

metonymization, a process by which a representamen is assigned a new referent



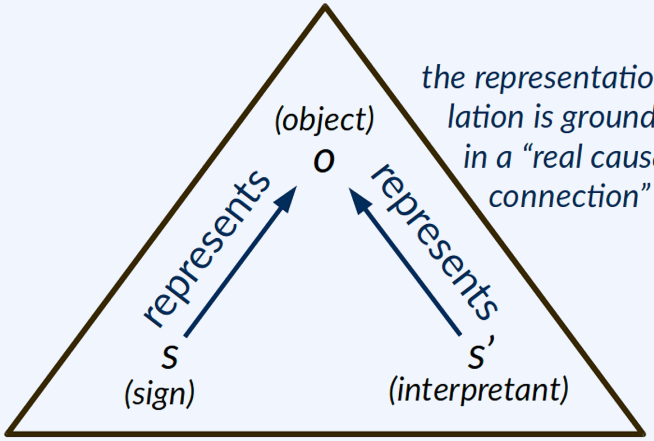
C. S. Peirce

Elementary Multi-perspective Material Ontology (EMMO)

The EMMO combines this with mereotopology – foundational ontology as **mereosemiotics**.

Cognitive process model

Peircean semiotics



the semiosis, a process by which a new representamen, the interpretant, is created

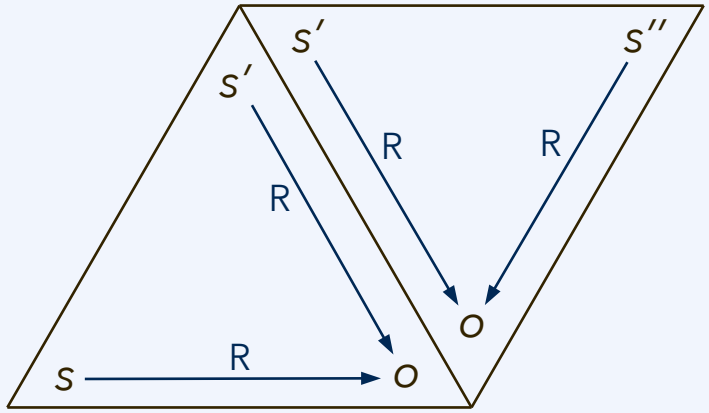
Each cognitive step starts from one representation relation, *e.g.*, *Rso*, and creates a new one, *Rs'o*.

The successor step reuses *Rs'o* and creates the next relation, *Rs''o*.

Cognitive process (example):

- First, experimental data *s* for material *o* are used to parameterize a model, obtaining model *s'*.
- Then, a simulation is done using model *s'*, yielding the simulation result *s''* (which also represents *o*).

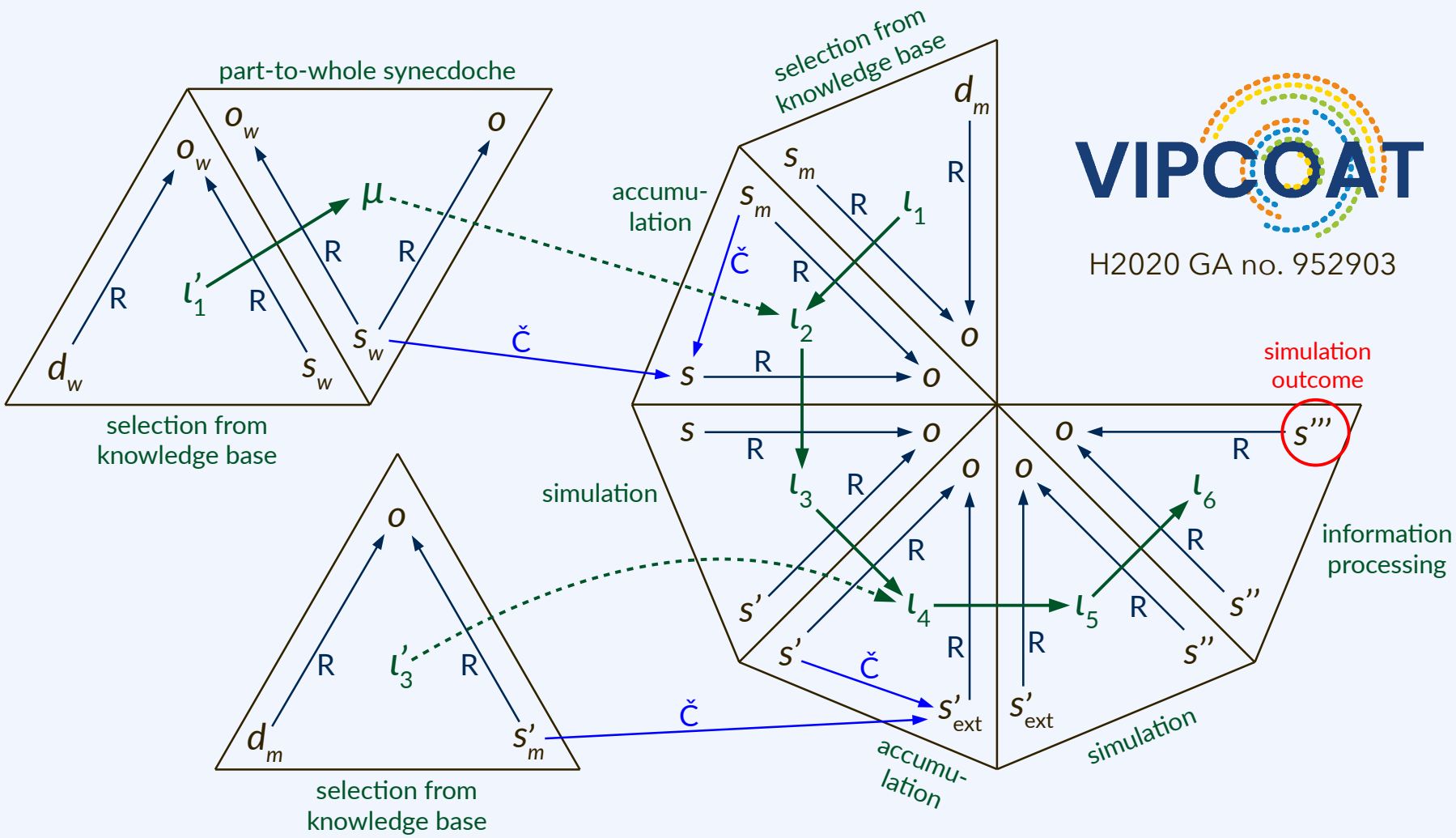
Research workflows as cognitive processes:



"is representamen for" is here abbreviated by R

Here, the first semiosis **directly grounds** the second semiosis.

Workflows as cognitive processes[1]

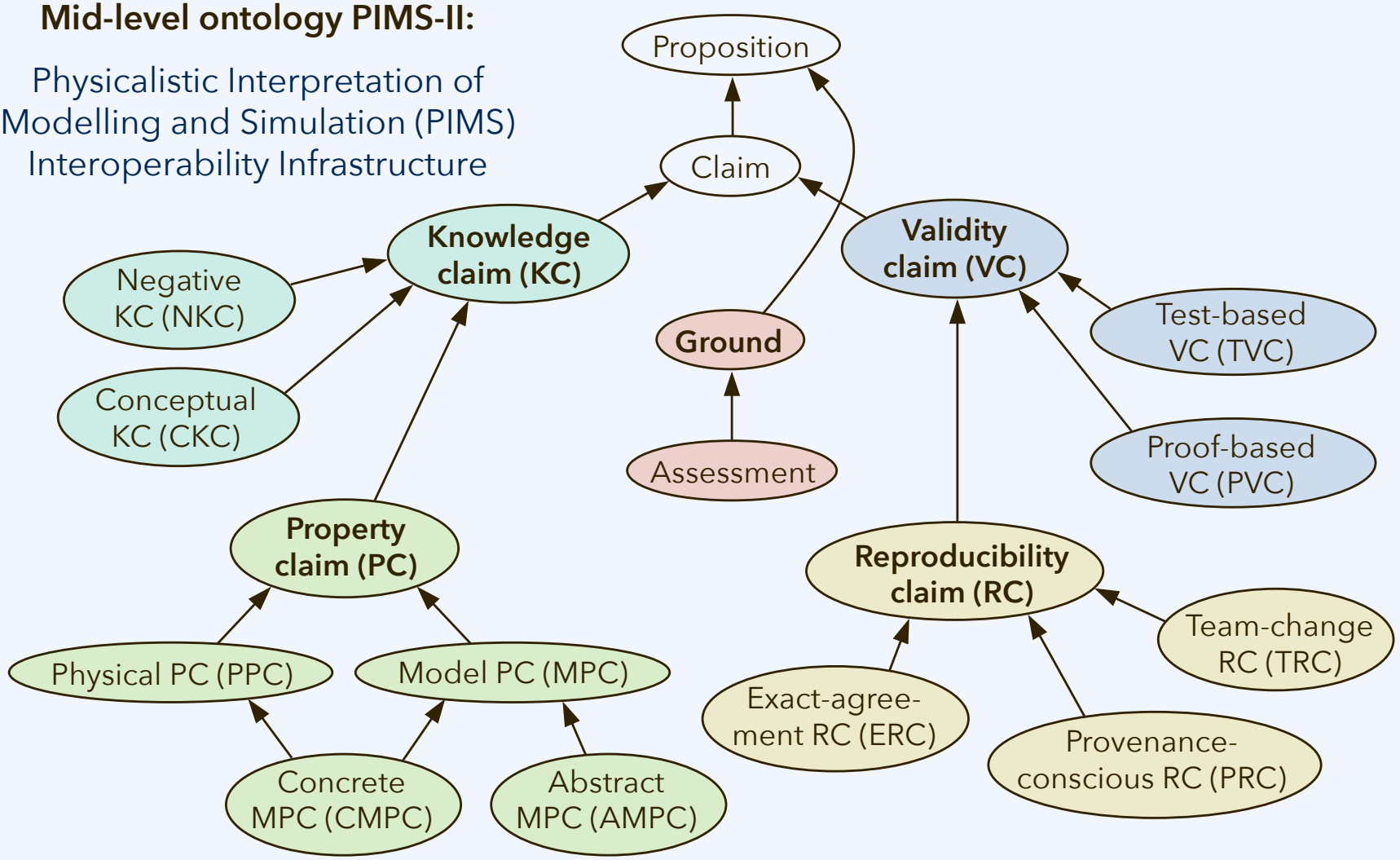


[1]P. Klein *et al.*, no. 26 in *Proc. JOWO 2021*, **2021**.

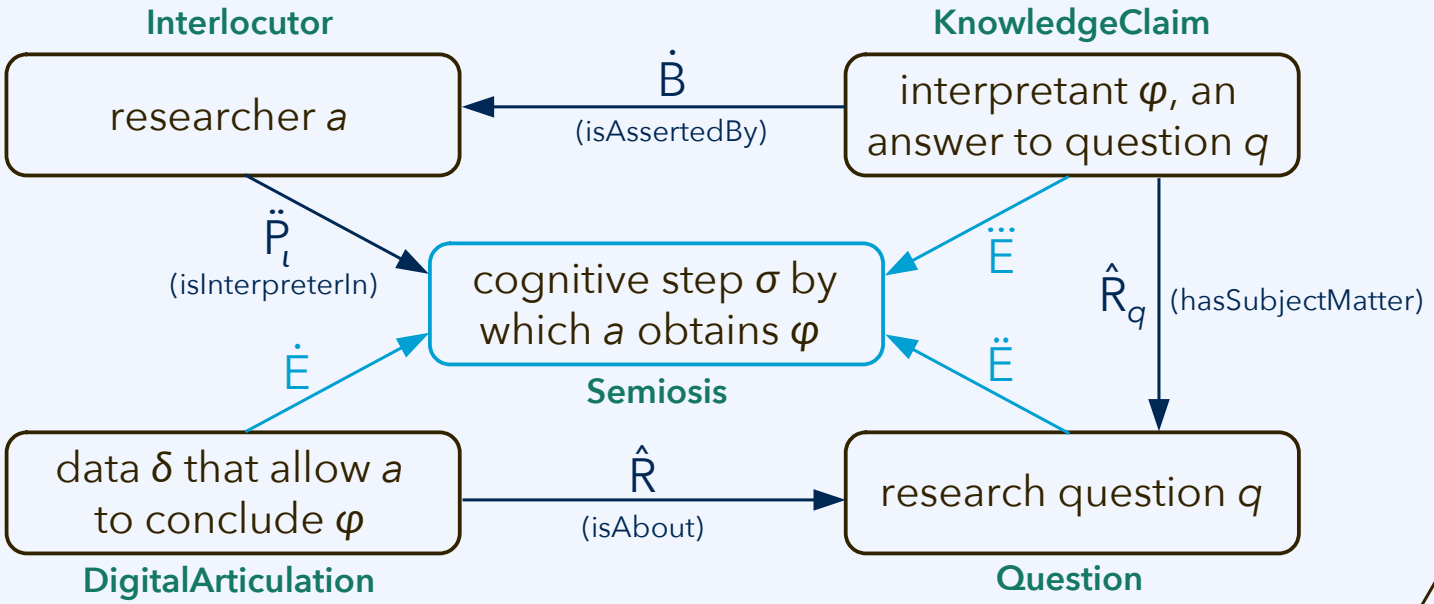
Knowledge and validity claims

Mid-level ontology PIMS-II:

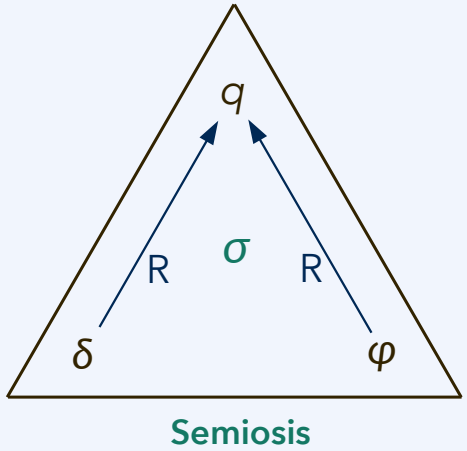
Physicalistic Interpretation of Modelling and Simulation (PIMS) Interoperability Infrastructure



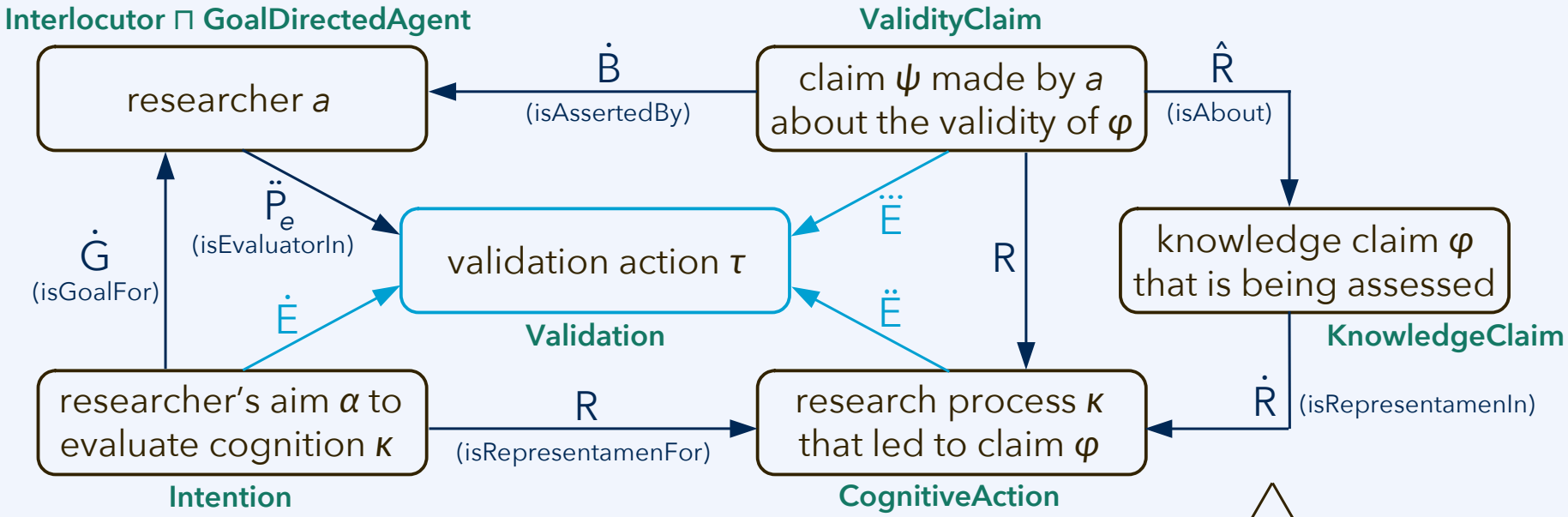
Knowledge claim schema



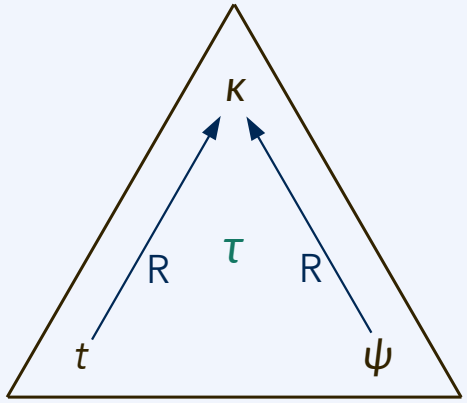
- The data are about the research problem, hence δ is a representamen for q; it has the role of the **sign**.
- As an outcome, a claim φ is obtained, which is a new representamen: The **interpretant**.



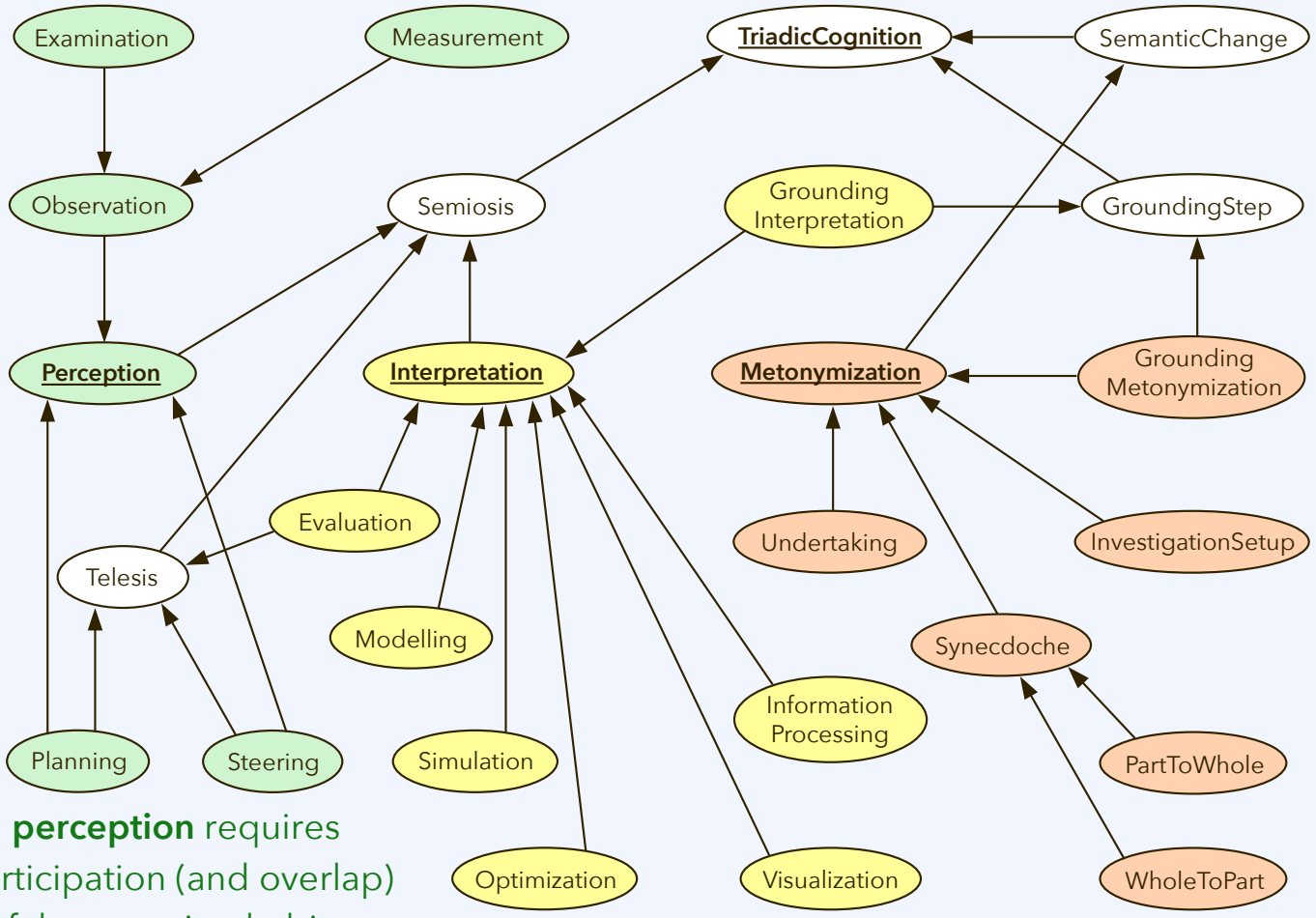
Validity claim schema



A validation is an evaluation where the evaluated object is a cognitive action and the interpretant is a validity claim.



Cognitive steps: Taxonomy



perception requires participation (and overlap) of the perceived object

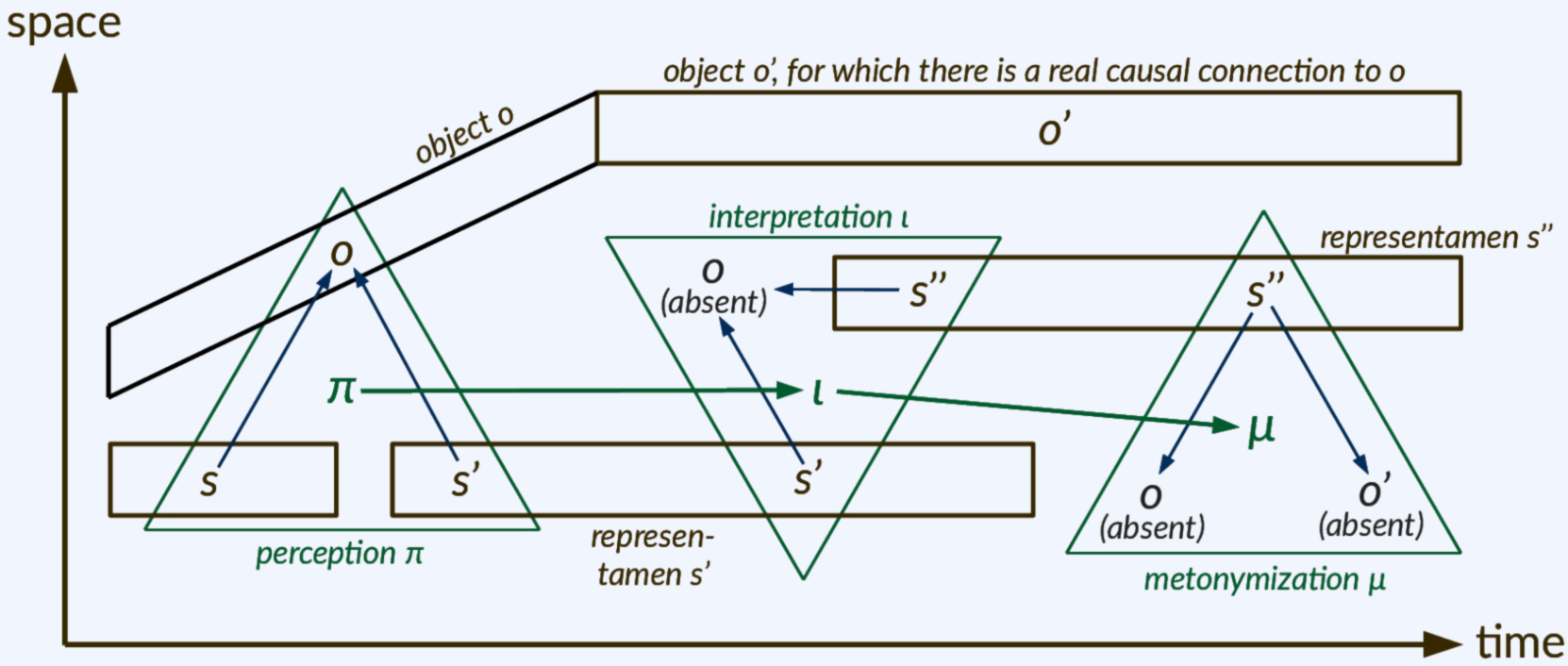
interpretation and **metonymization** do not entail physical participation of the referents

metonymization preserves the „real causal connection" (Peirce) between the sign and its old & new referents

Cognitive steps in mereosemiotics

PIMS-II mid-level ontology:[1, 2] http://www.molmod.info/semantics/pims-ii.ttl

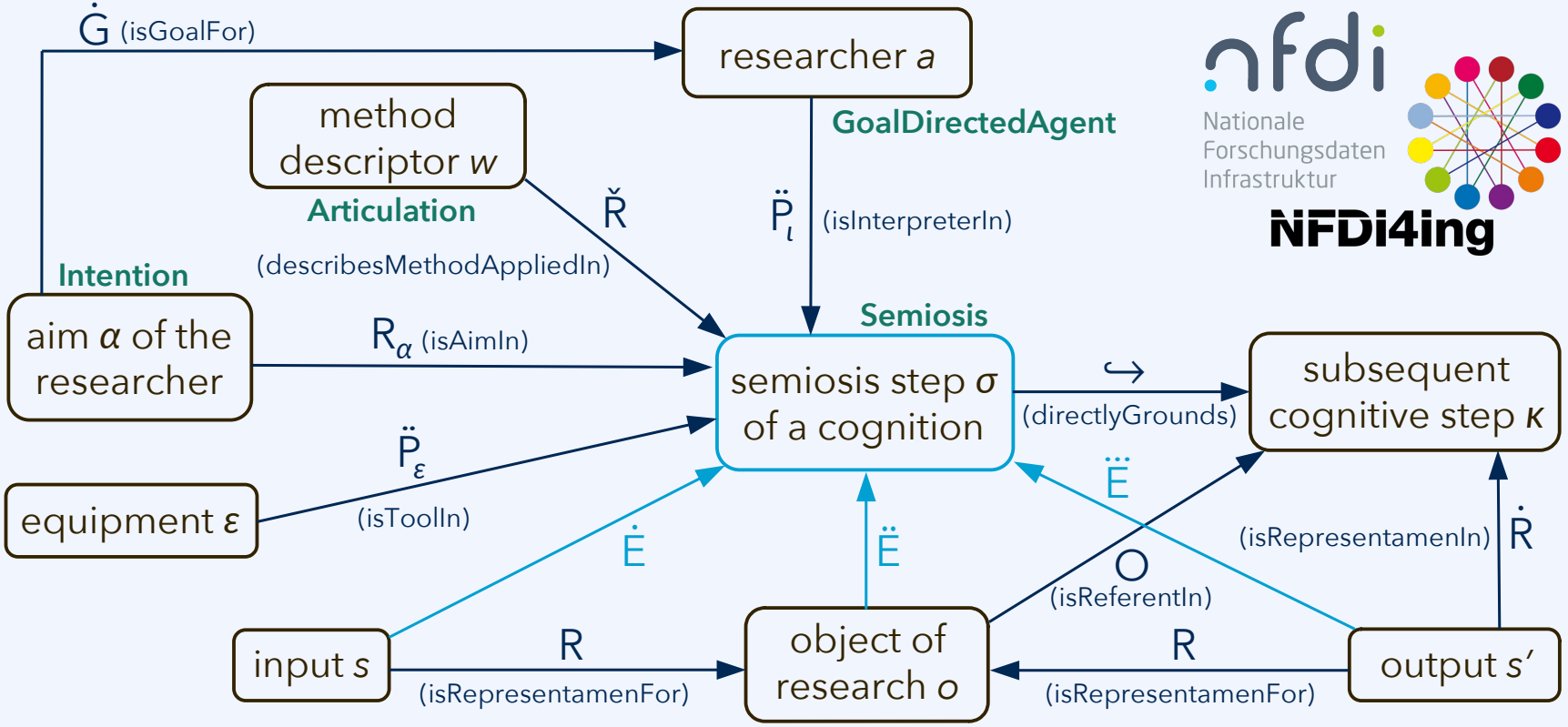
Mereosemiotics:[1-3] Combination of mereotopology and Peircean semiotics



[1]M. T. Horsch, no. 3 in *Proc. JOWO 2021*, **2021**. [2]P. Klein *et al.*, no. 26 in *Proc. JOWO 2021*, **2021**.
[3]M. T. Horsch, S. Chiacchiera, B. Schembera, M. Seaton, I. T. Todorov, in *Proc. ECCOMAS 2020*, **2021**.

Research process documentation

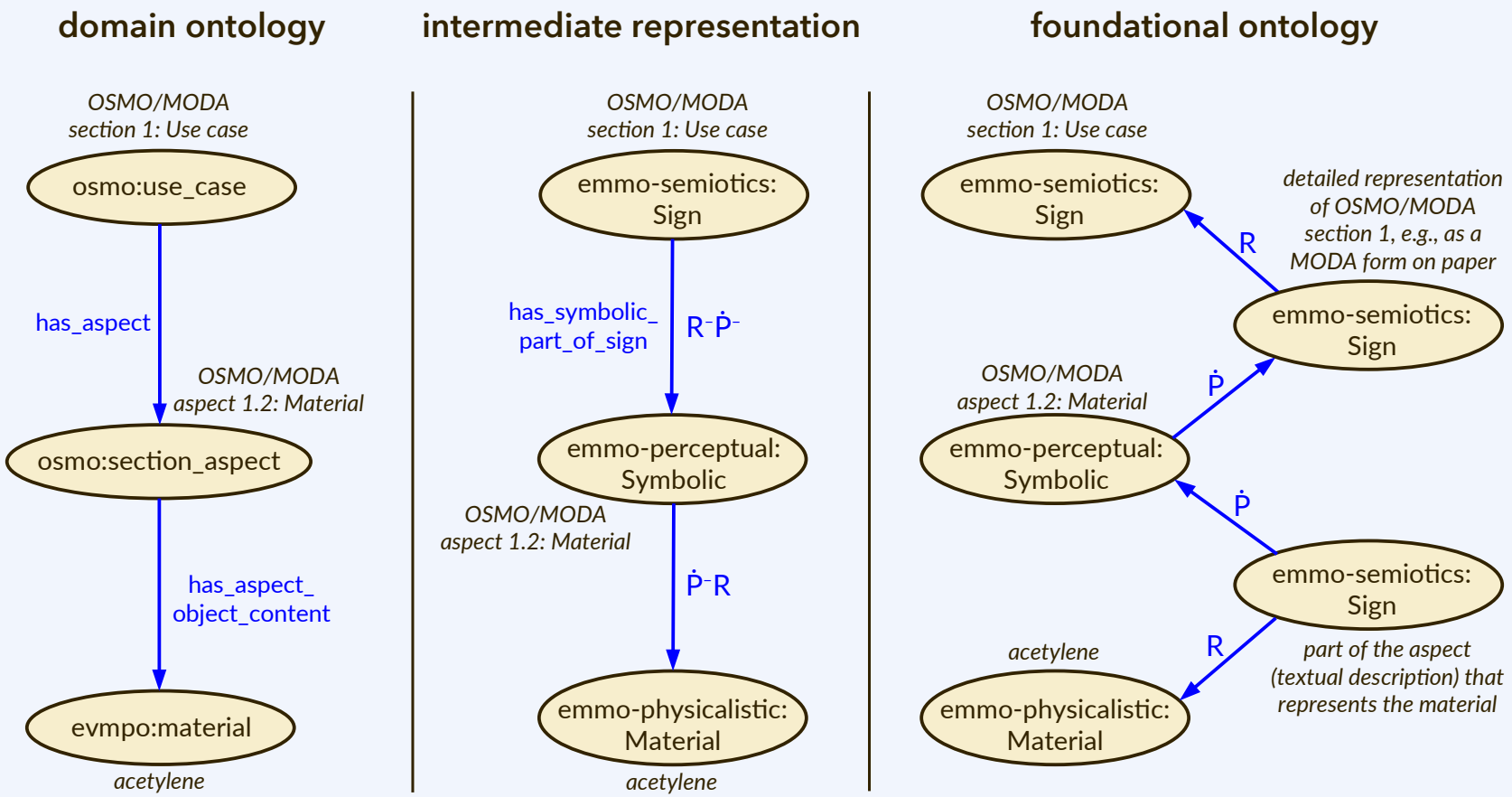
PIMS-II schema[1] for a **semiosis** step,[2] aligned with **processing step** from **m4i**.[3]



[1]M. T. Horsch, no. 3 in *Proc. JOWO 2021*, **2021**.
[2]M. T. Horsch, *Mereosemiotics: Five scenarios* (*cf.* Borgo *et al.*'s top-level ontology comparison), **2021**.
[3]https://w3id.org/nfdi4ing/metadata4ing/, **2022**.

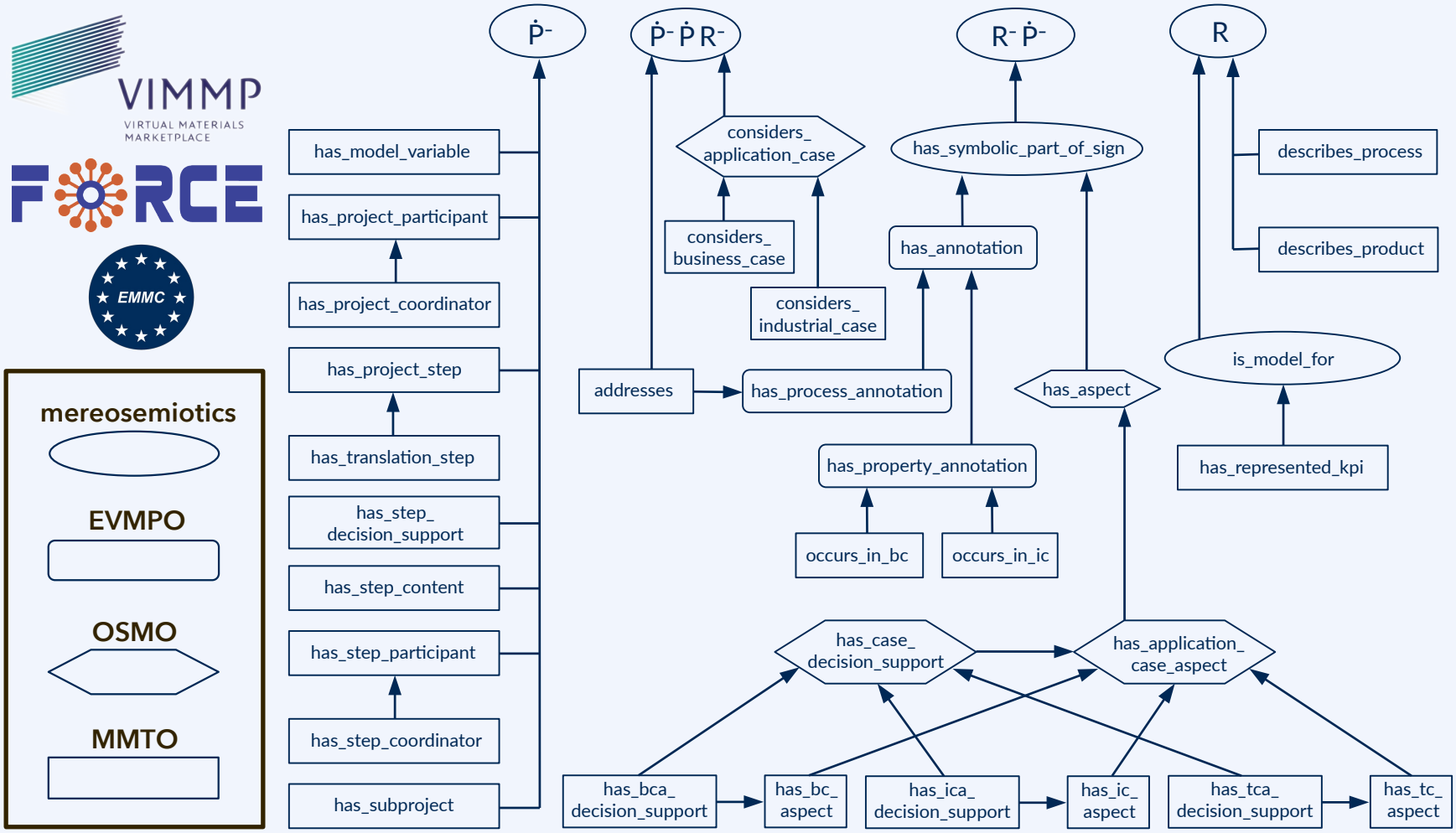
Mereosemiotic chain relations

Ontology alignment work from H2020 project VIMMP:[1]



[1]M. Horsch, S. Chiacchiera, W. Cavalcanti, B. Schembera, *Data Technology in Materials Modelling*, Springer, **2021**.

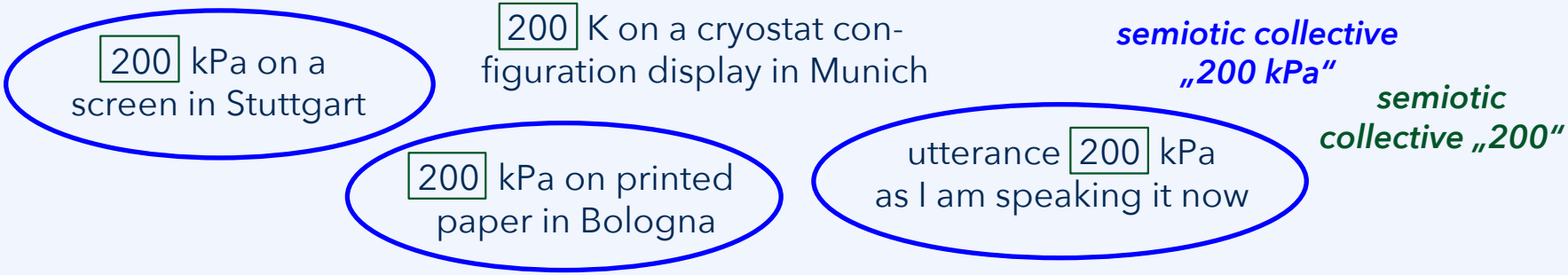
Mereosemiotic chain relations[1]



[1]M. T. Horsch *et al.*, *pp.* 45–59 in *Proc. DAMDID 2020*, Springer, CCIS no. **1427**, doi:10.1007/978-3-030-81200-3_4, **2021**.

Semiotic collectives

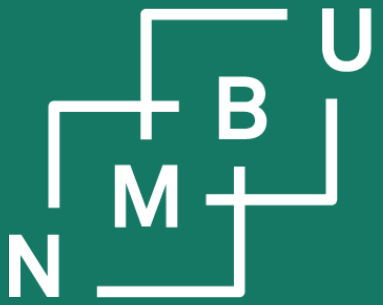
EMMO is based on nominalism with objects defined by 4D spacetime regions. This creates some challenges[1] when working with data; *e.g.*, take "200 kPa":



Another example: „The stadiums of the Premier League have a total capacity of 833 000 people." The referent of the property are all the stadiums together, as a collective; none of the individual stadiums holds 833 000 people.

For such purposee, PIMS-II defines **semiotic collectives**.

[1]M. Horsch, S. Chiacchiera, B. Schembera, M. Seaton, I. Todorov, *Proc. WCCM-ECCOMAS 2020*, **2021**.



Norges miljø- og biovitenskapelige universitet

Documentation of epistemic metadata by a mid-level ontology of cognitive processes

Martin Thomas Horsch[1] and Björn Schembera[2]

[1]Norwegian University of Life Sciences, Department of Data Science
[2]University of Stuttgart, Institute of Applied Analysis and Numerical Simulation

http://www.molmod.info/semantics/pims-ii.ttl

CAOS VI @ JOWO 2022 18th August 2022 Jönköping