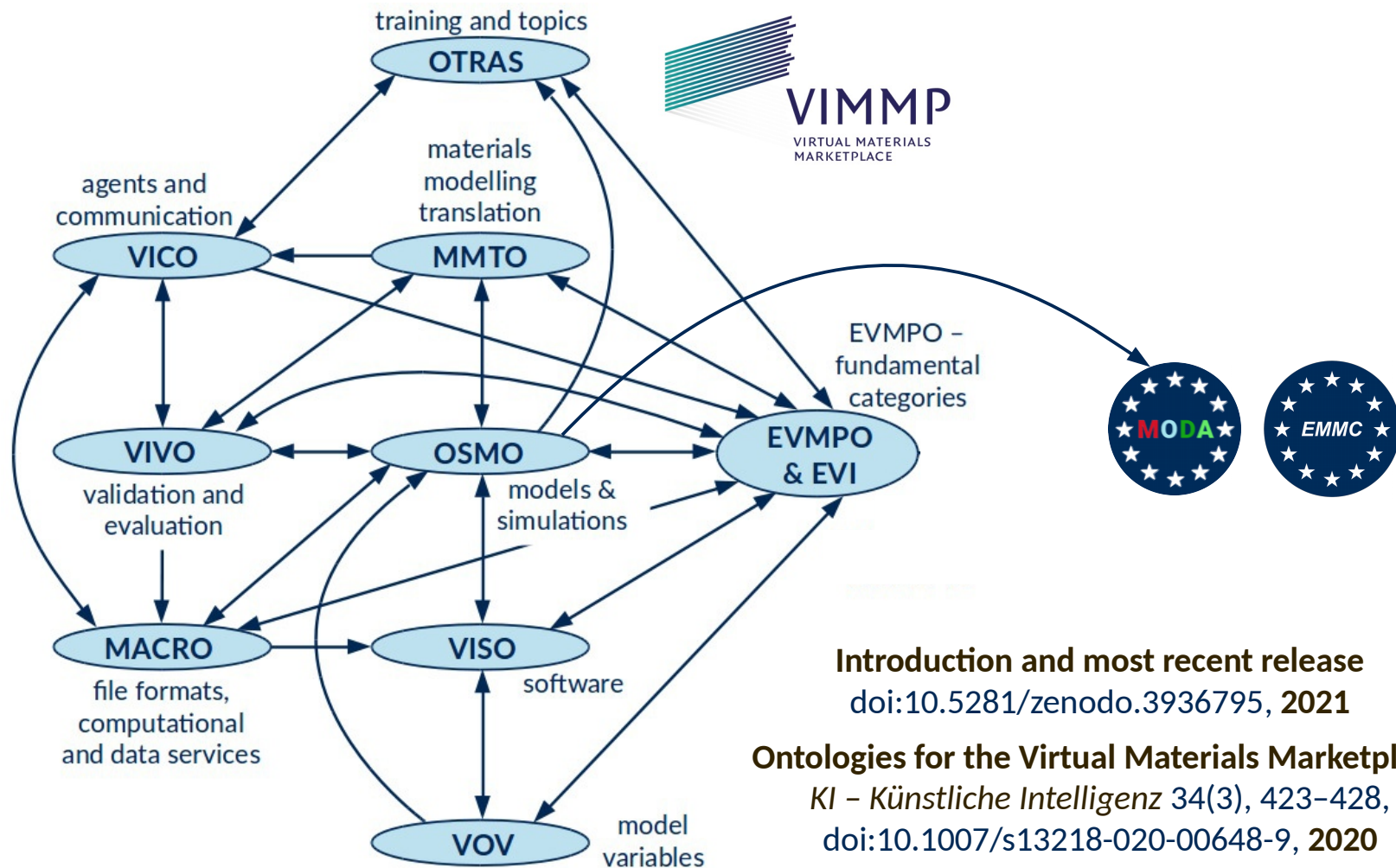# Knowledge representation for materials modelling marketplaces

M. Horsch,[1, 2] S. Chiacchiera,[2] M. Seaton,[2] I. Todorov,[2] P. Klein,[3] N. Konchakova,[4] C. Niethammer,[1] J. Vrabec,[5] D. Toti,[6, 7] G. Mogni,[7] G. Goldbeck,[7] P. Schiffels,[8] and W. Cavalcanti,[8] *High Performance Computing Center Stuttgart,[1] UKRI STFC Daresbury Laboratory,[2] Fraunhofer ITWM,[3] Helmholtz-Zentrum Hereon,[4] Techn. Univ. Berlin,[5] Catholic Univ. Brescia,[6] Goldbeck Consulting Ltd.,[7] Fraunhofer IFAM[8]*
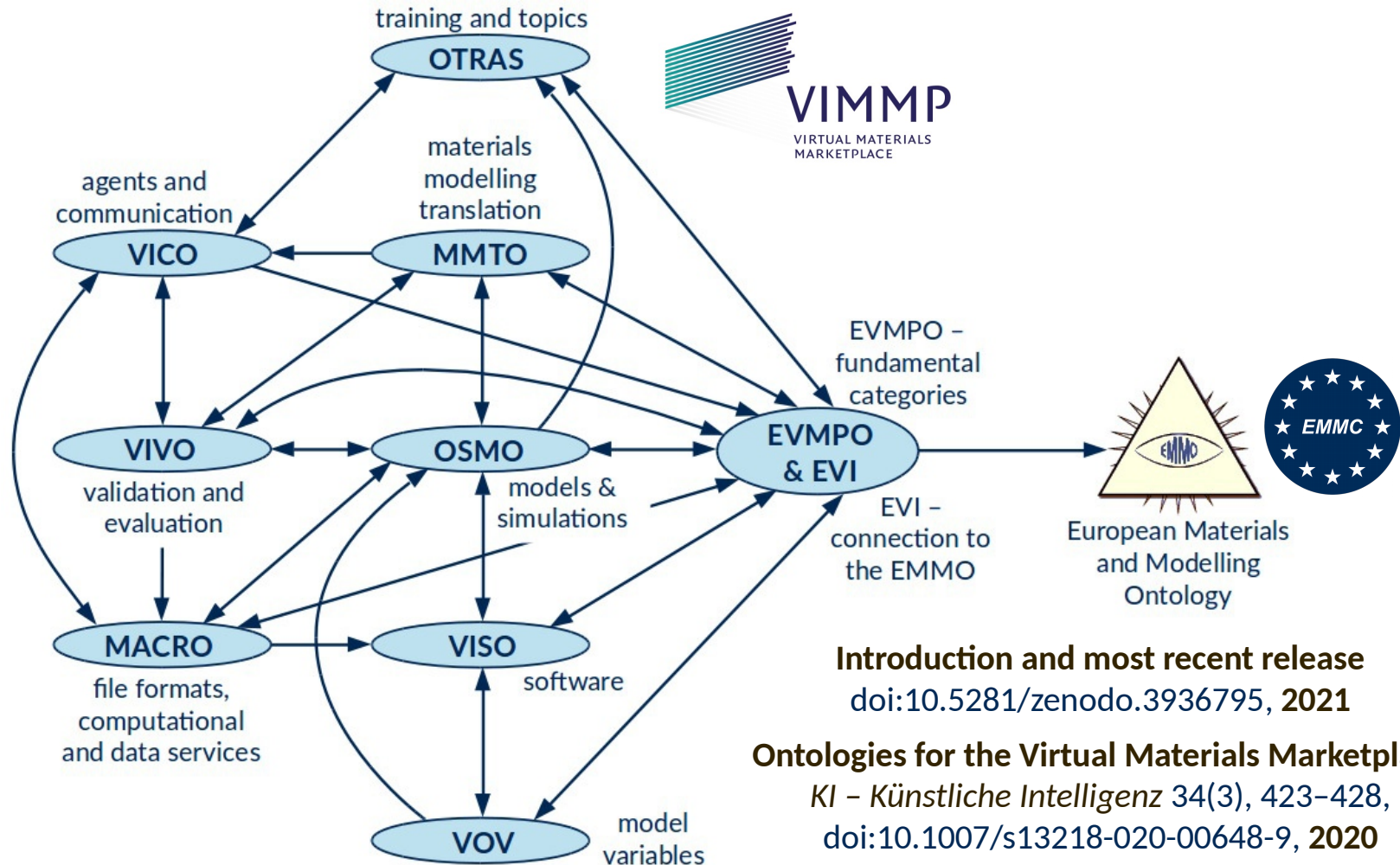
EMMC-guided ontology design for materials modelling

Introduction and most recent release
doi:10.5281/zenodo.3936795, **2021**

Ontologies for the Virtual Materials Marketplace
KI – Künstliche Intelligenz 34(3), 423–428,
doi:10.1007/s13218-020-00648-9, **2020**

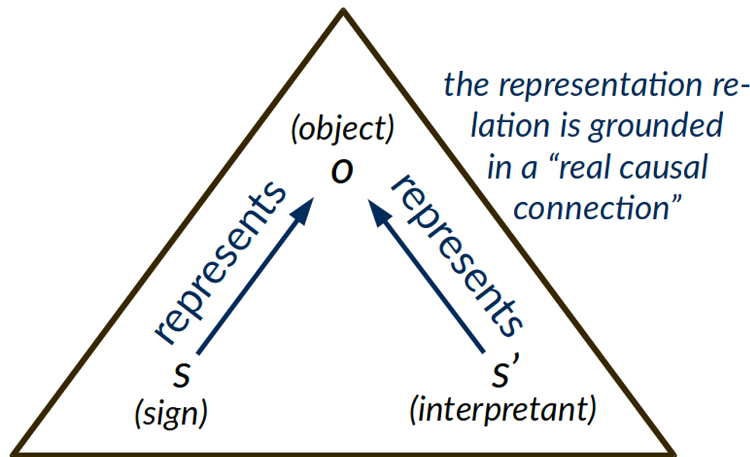# EMMO-based ontology design for materials modelling

Introduction and most recent release
doi:10.5281/zenodo.3936795, **2021**

Ontologies for the Virtual Materials Marketplace
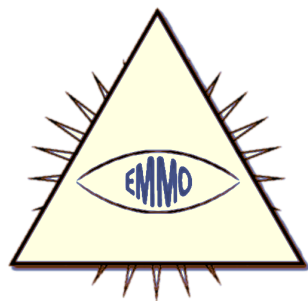*KI – Künstliche Intelligenz* 34(3), 423–428,
doi:10.1007/s13218-020-00648-9, **2020**

EMMO-based ontology design for materials modelling

Peircean semiotics

(object)
O

represents | represents

S
(sign)

S'
(interpretant)

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

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

C. S. Peirce

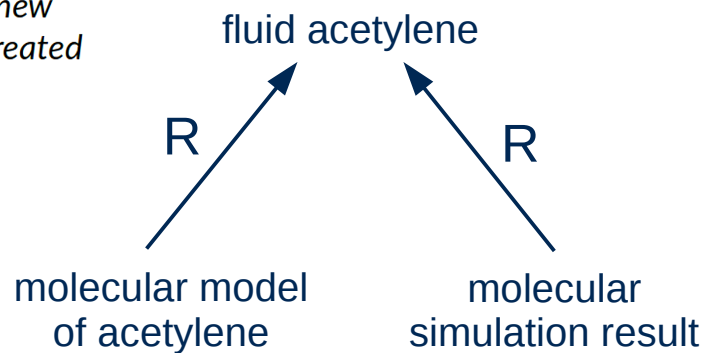## European Materials and Modelling Ontology[1]

1) Taxonomy:
Conceptual hierarchy (subclass relation)

2) Semiotics:
Representation of physical entities by signs

3) Mereotopology:
Spatiotemporal parthood and connectivity

fluid acetylene

R   R

molecular model of acetylene

molecular simulation result

"represents" or **"is sign for"** is here abbreviated by **R**

[1]G. Goldbeck *et al.*, *Proc. NAFEMS World Congress*, NWC_19_86, **2019**.

EMMO-based ontology design for materials modelling

electronic

atomistic

mesoscopic

continuum

European Materials and Modelling Ontology[1]

1) **Taxonomy**:
   Conceptual hierarchy (subclass relation)

2) **Semiotics**:
   Representation of physical entities by signs

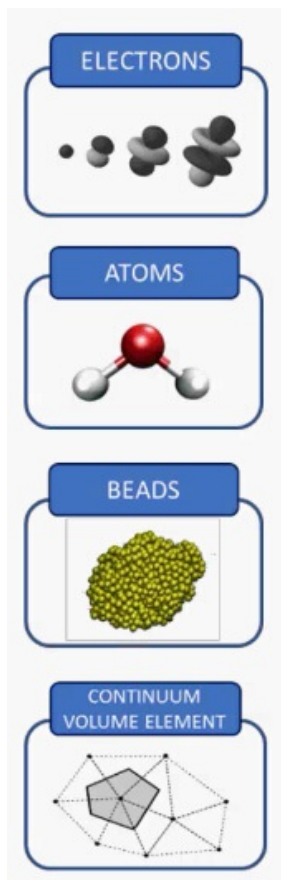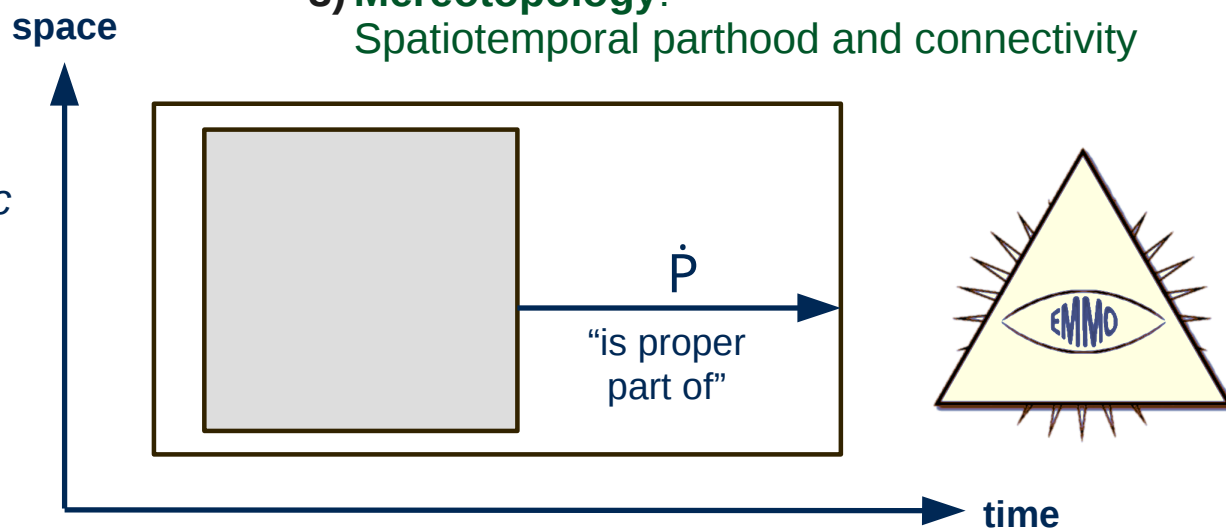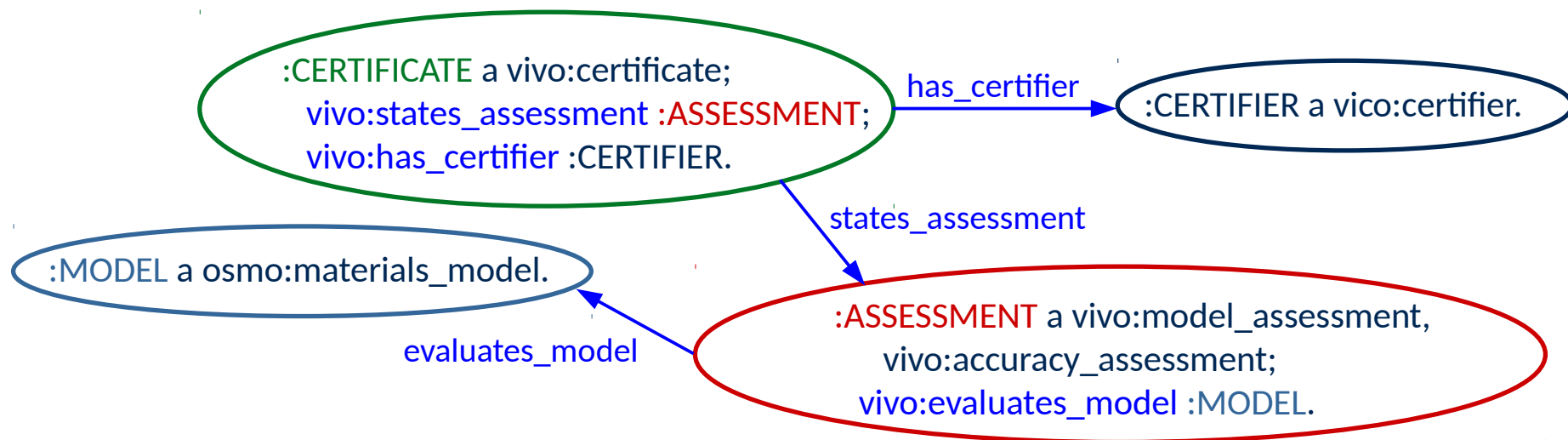3) **Mereotopology**:
   Spatiotemporal parthood and connectivity
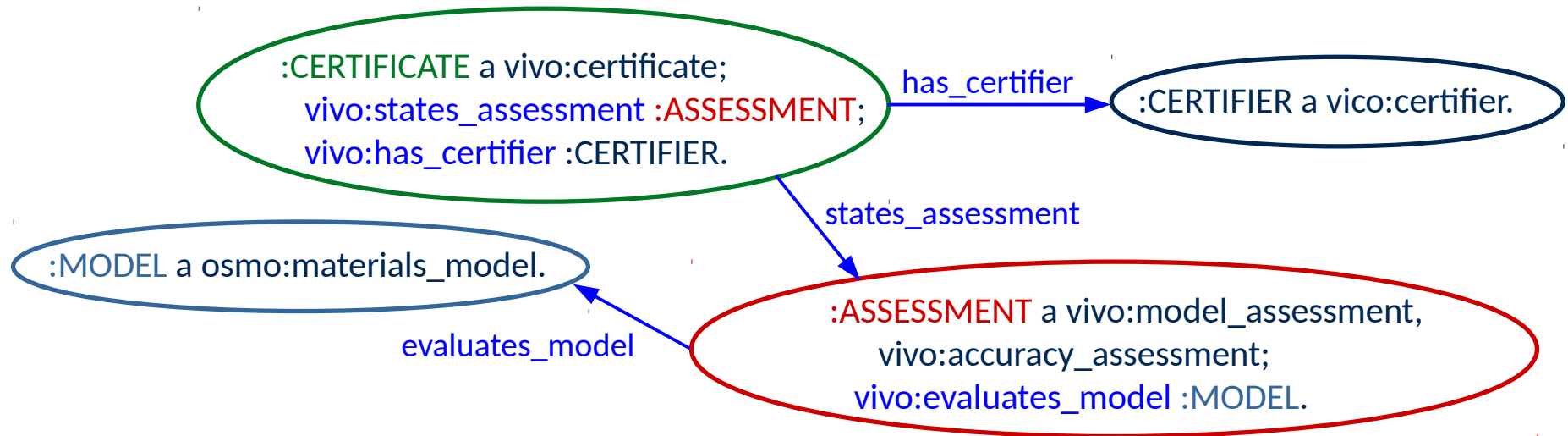

space

$\dot{P}$

"is proper part of"

time

[1]G. Goldbeck *et al.*, *Proc. NAFEMS World Congress*, NWC_19_86, **2019**.

Knowledge graphs instantiating the VIMMP ontologies

:CERTIFICATE a vivo:certificate;
  vivo:states_assessment :ASSESSMENT;
  vivo:has_certifier :CERTIFIER.

has_certifier →

:CERTIFIER a vico:certifier.

states_assessment

:MODEL a osmo:materials_model.

evaluates_model

:ASSESSMENT a vivo:model_assessment,
  vivo:accuracy_assessment;
  vivo:evaluates_model :MODEL.

The certifier :CERTIFIER has issued a certificate (the IRI of which is :CERTIFICATE) stating a model accuracy assessment (the IRI of which is :ASSESSMENT) that evaluates the materials model :MODEL.

Knowledge graphs instantiating the VIMMP ontologies

:CERTIFICATE a vivo:certificate;
  vivo:states_assessment :ASSESSMENT;
  vivo:has_certifier :CERTIFIER.

has_certifier → :CERTIFIER a vico:certifier.

states_assessment

:MODEL a osmo:materials_model.

:ASSESSMENT a vivo:model_assessment,
  vivo:accuracy_assessment;
  vivo:evaluates_model :MODEL.

evaluates_model

The certifier :CERTIFIER has issued a certificate (the IRI of which is :CERTIFICATE) stating a model accuracy assessment (the IRI of which is :ASSESSMENT) that evaluates the materials model :MODEL.
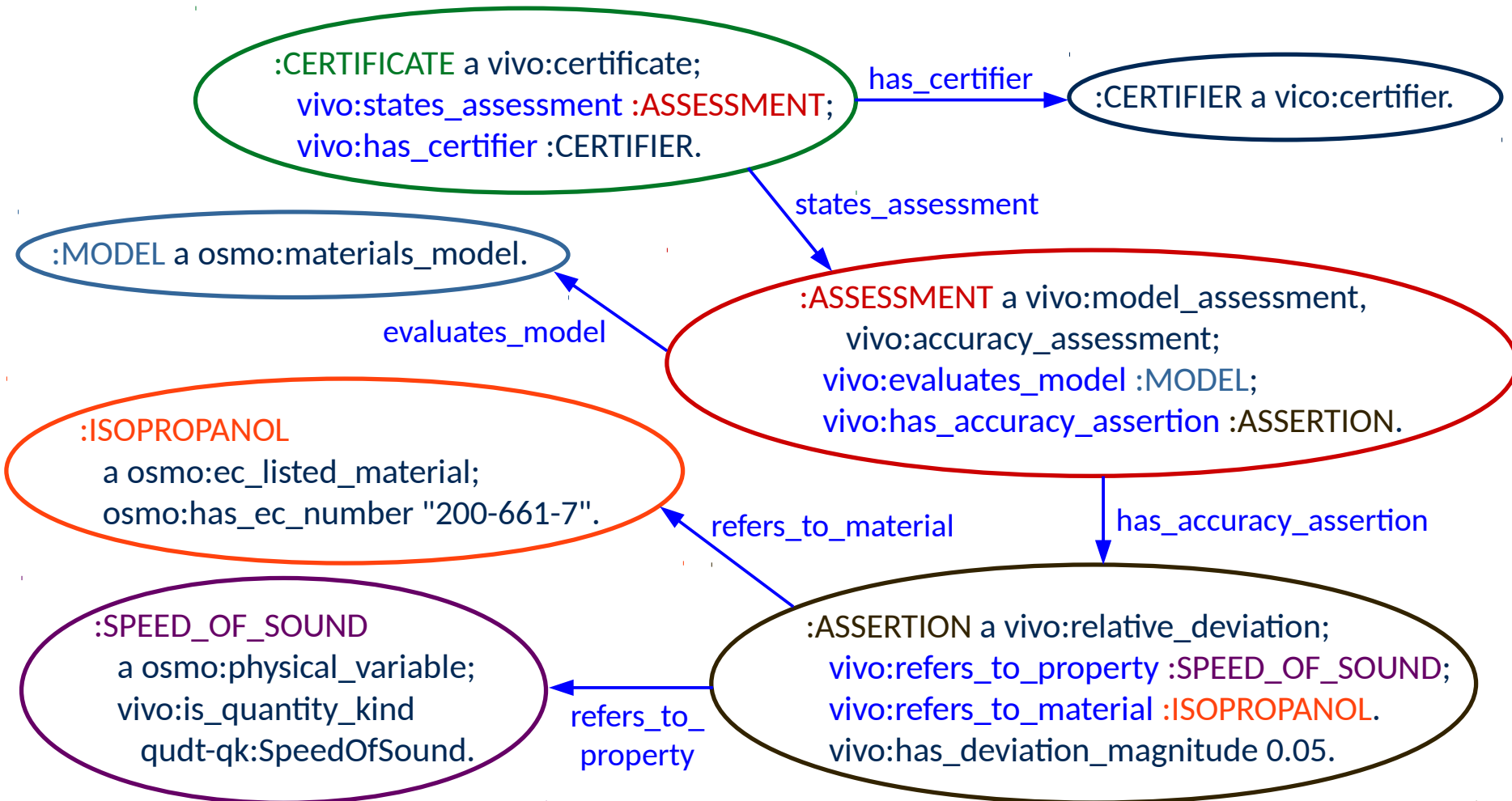
In the model accuracy assessment :ASSESSMENT, it is asserted that the materials model :MODEL has a relative error of 5% for the speed of sound of isopropanol.

Knowledge graphs instantiating the VIMMP ontologies

:CERTIFICATE a vivo:certificate;
vivo:states_assessment :ASSESSMENT;
vivo:has_certifier :CERTIFIER.

has_certifier → :CERTIFIER a vico:certifier.

states_assessment

:MODEL a osmo:materials_model.

evaluates_model

:ASSESSMENT a vivo:model_assessment,
vivo:accuracy_assessment;
vivo:evaluates_model :MODEL;
vivo:has_accuracy_assertion :ASSERTION.

:ISOPROPANOL
a osmo:ec_listed_material;
osmo:has_ec_number "200-661-7".

refers_to_material

has_accuracy_assertion

:SPEED_OF_SOUND
a osmo:physical_variable;
vivo:is_quantity_kind
qudt-qk:SpeedOfSound.

refers_to_property

:ASSERTION a vivo:relative_deviation;
vivo:refers_to_property :SPEED_OF_SOUND;
vivo:refers_to_material :ISOPROPANOL.
vivo:has_deviation_magnitude 0.05.

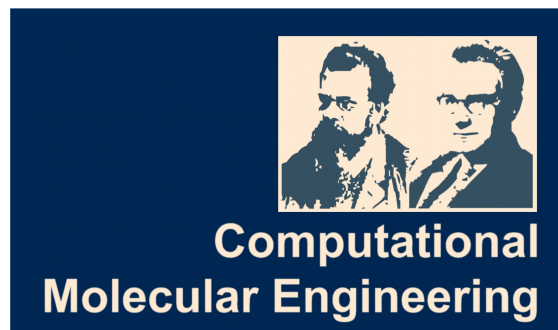# Application to the MolMod DB molecular model database

Geometry

Types and positions
of interaction sites

Dispersion and repulsion

Lennard-Jones or Mie potential:
Size and energy parameters

Electrostatics

Point charge or multipole
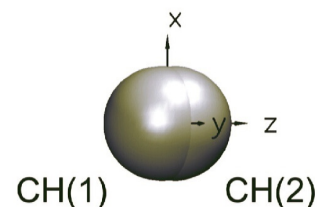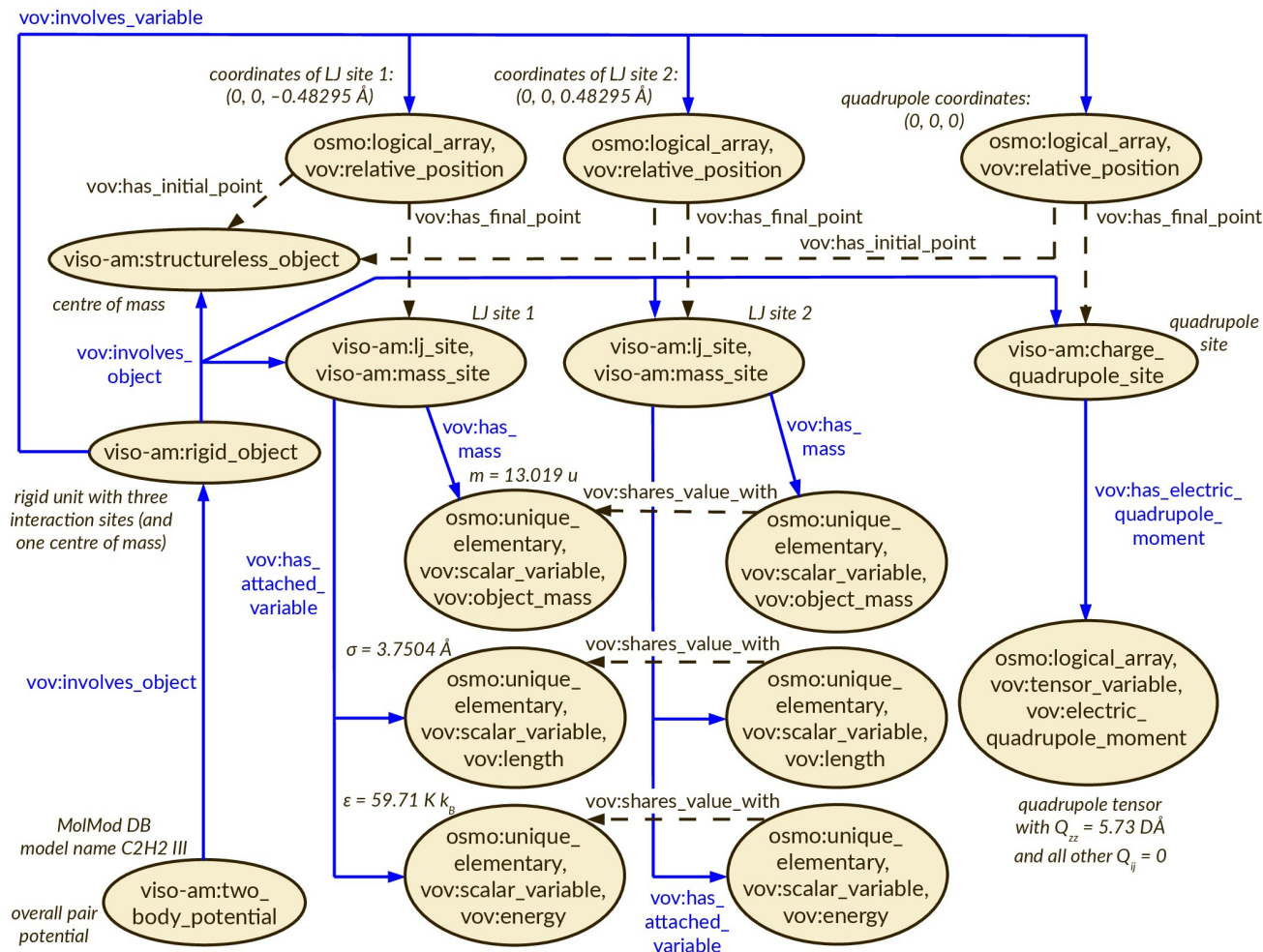(point dipole or quadrupole):
Magnitude and orientation

Molecular model database (MolMod DB)

http://molmod.boltzmann-zuse.de/

Pair potentials for over 150 molecular fluids

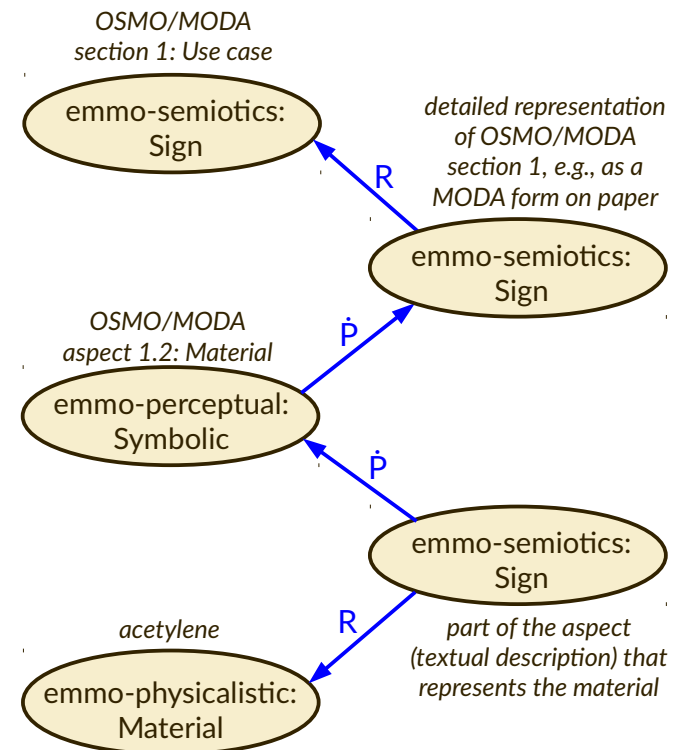# Application to the MolMod DB molecular model database

Ontology alignment („crosswalks")

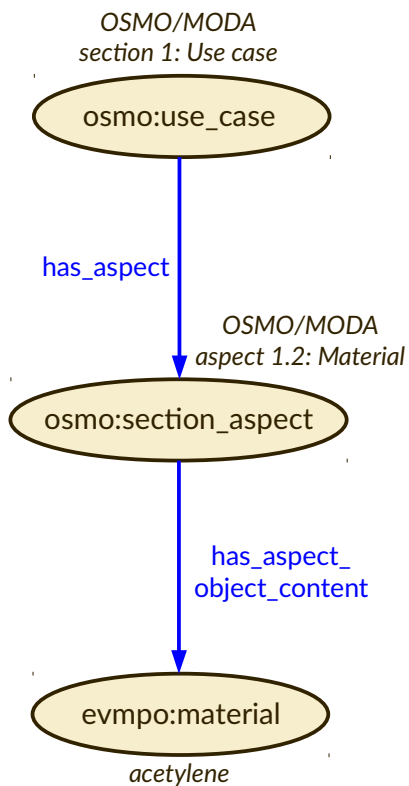**marketplace-level domain ontology representation**

top-level foundational ontology representation with unfolded chain relations
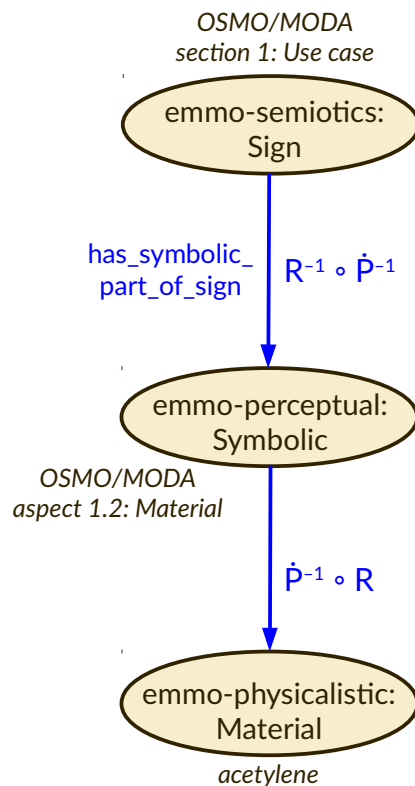

[1]M. T. Horsch, S. Chiacchiera, W. L. Cavalcanti, B. Schembera, *Data Technology in Materials Modelling*, Springer, **2021**.
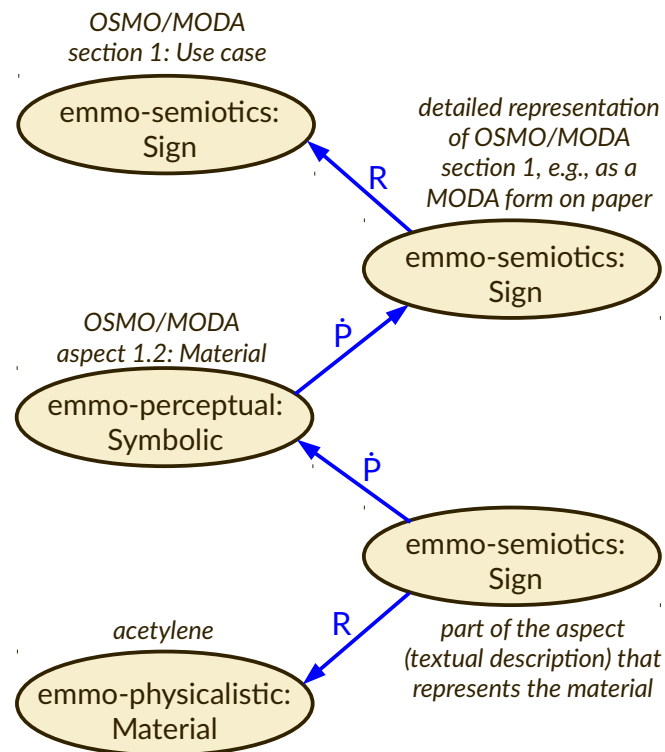
Ontology alignment („crosswalks")

marketplace-level domain ontology representation

intermediate representation using mereosemiotic chain relations

top-level foundational ontology representation with unfolded chain relations

OSMO/MODA section 1: Use case

osmo:use_case

has_aspect

OSMO/MODA aspect 1.2: Material

osmo:section_aspect

has_aspect_object_content

evmpo:material

acetylene

OSMO/MODA section 1: Use case

emmo-semiotics: Sign

has_symbolic_part_of_sign $R^{-1} \circ \dot{P}^{-1}$

emmo-perceptual: Symbolic

OSMO/MODA aspect 1.2: Material

$\dot{P}^{-1} \circ R$

emmo-physicalistic: Material

acetylene

OSMO/MODA section 1: Use case

emmo-semiotics: Sign

detailed representation of OSMO/MODA section 1, e.g., as a MODA form on paper

R

emmo-semiotics: Sign

$\dot{P}$

OSMO/MODA aspect 1.2: Material

emmo-perceptual: Symbolic

$\dot{P}$

emmo-semiotics: Sign

acetylene

R

emmo-physicalistic: Material

part of the aspect (textual description) that represents the material

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

Digital infrastructures and communication of knowledge

At the **Battle of Hattin**, 20 000 valiant crusaders sacrificed their lives in defense of the Holy Land.

„*Saladin ordered that they should be beheaded, choosing to have them dead rather than in prison.* **With him was a whole band of scholars**" (Imad ed-Din, secretary of Saladin)
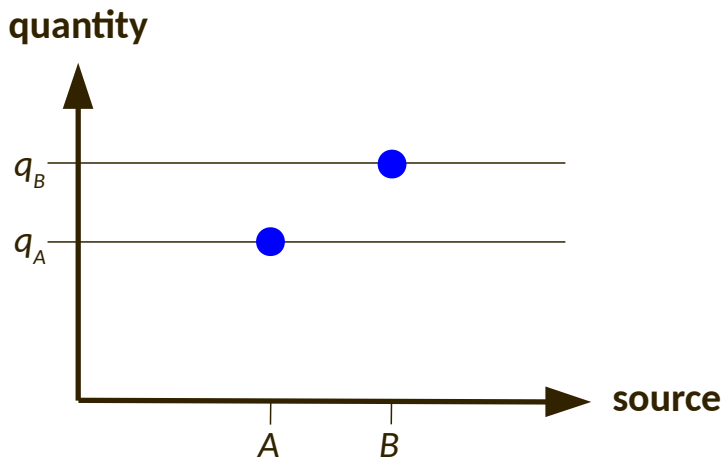
Battle of Hattin (1187)

Saracens

Averroes[1]
(Bidayat al-Mujtahid)
al-Sulami
(Book of the Jihad, 1106)

Crusaders

Pope Urban II
(in *Proc. Council Clermont*, 1059)
Augustine (*Civitas Dei*, 426)

[1]Research data infrastructure on Averroes' works: https://averroes.uni-koeln.de/

Digital infrastructures and communication of knowledge

quantity

q_B

q_A

source

A   B

Saracens   **Crusaders**

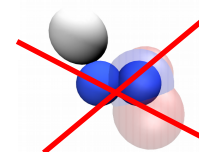Averroes[1]
(Bidayat al-Mujtahid)
al-Sulami
(Book of the Jihad, 1106)

Pope Urban II
(in *Proc. Council Clermont*, 1059)
Augustine (*Civitas Dei*, 426)

[1]Research data infrastructure on Averroes' works: https://averroes.uni-koeln.de/

Digital infrastructures and communication of knowledge
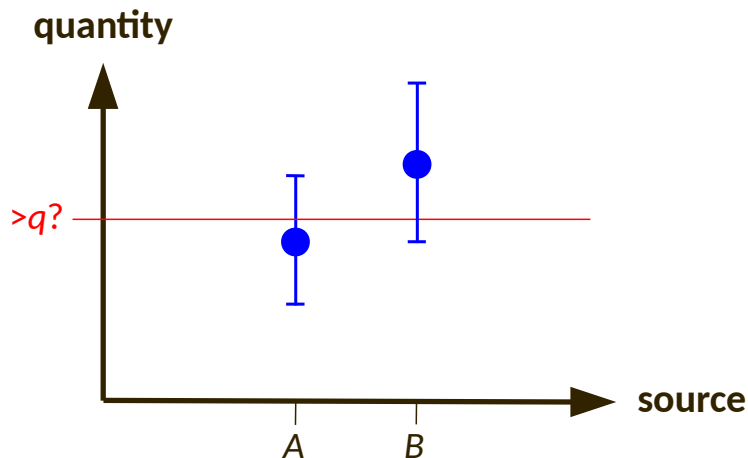
quantity

>q?

A   B   source

Saracens   Crusaders

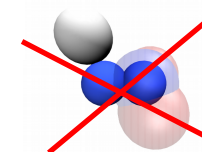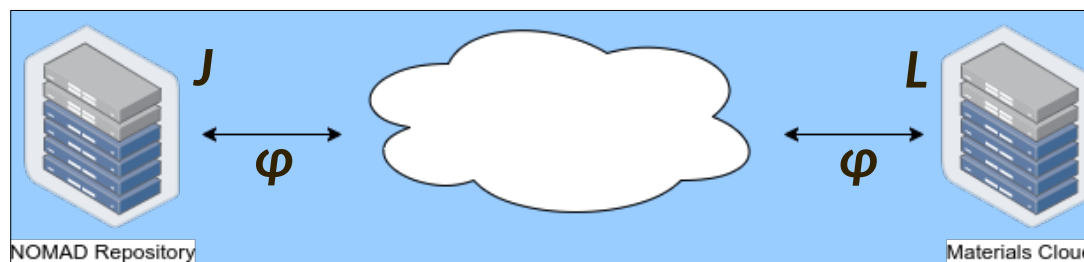Averroes[1]
(Bidayat al-Mujtahid)
al-Sulami
(Book of the Jihad, 1106)

Pope Urban II
(in *Proc. Council Clermont*, 1059)
Augustine (*Civitas Dei*, 426)

[1]Research data infrastructure on Averroes' works: https://averroes.uni-koeln.de/

Digital infrastructures and communication of knowledge

Digital platforms and infrastructures can store and exchange scientific knowledge.
Scientific knowledge is a kind of knowledge (or little will qualify as knowledge).

Scenario requiring epistemological formalization:

- "The scientific knowledge φ is communicated by knowledge base J to L."

- It would be inappropriate to require every φ to be a justified true belief.

Digital infrastructures and communication of knowledge

Digital platforms and infrastructures can store and exchange scientific knowledge.
Scientific knowledge is a kind of knowledge (or little will qualify as knowledge).

Scenario requiring epistemological formalization:

- "The scientific knowledge φ is communicated by knowledge base J to L."

- It would be inappropriate to require every φ to be a justified true belief.
  Instead, φ is a justified tenable assertion, by disciplinary standards.

Digital infrastructures and communication of knowledge

Digital platforms and infrastructures can store and exchange scientific knowledge. Scientific knowledge is a kind of knowledge (or little will qualify as knowledge).
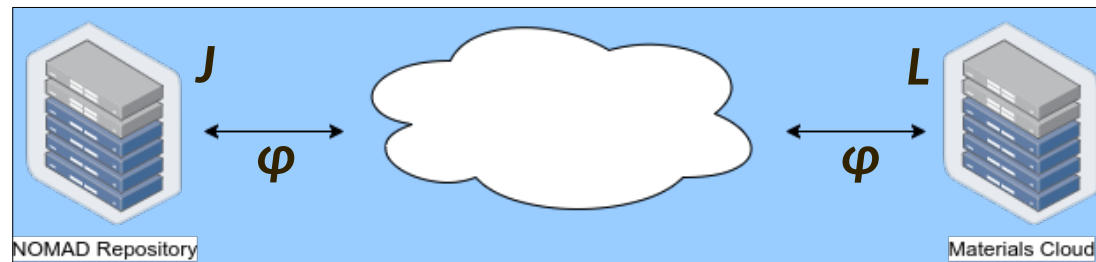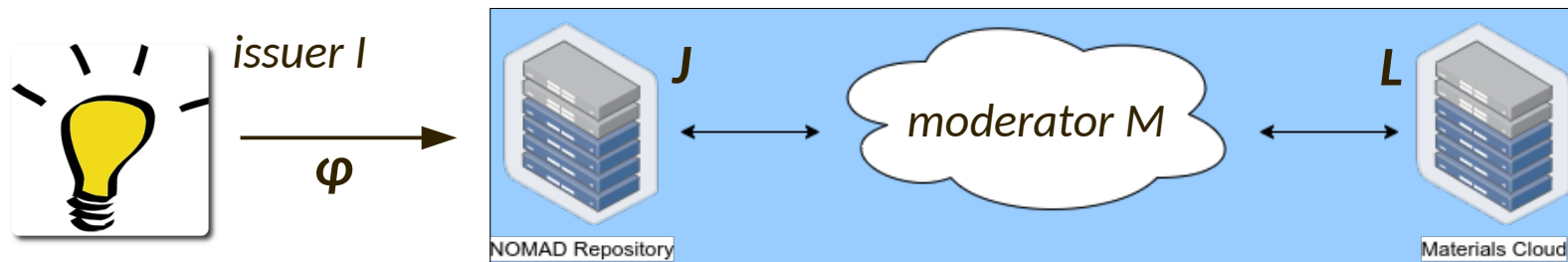

Scenario requiring epistemological formalization:

- "M asserts and approves $\varphi'(I, J, L, \varphi)$," where $\varphi'(I, J, L, \varphi)$ is given by:
- "The scientific knowledge φ, previously issued by a source I, has been communicated by the knowledge base J to the knowledge base L."
- J, L, and M have a justified true belief in φ'.
- φ is a justified tenable assertion, by the standards applied to I by M.

The challenge of „epistemic opacity"

Issue raised by Humphreys:[1] **Justification of φ** appears (to some) to be **opaque**.[1,2]

validation becomes „holistic"

alleged sources of opacity

no explicit justification is computed

stochastic elements (Monte Carlo)

Lyapunov instability (molecular dynamics)

[1]Knowledge of „*all epistemically relevant elements*" cannot be attained (Humphreys, 2004, 2011).
[2]Durán and Formanek (2018): „*epistemically relevant elements*" = „*steps of the […] justification*".

The challenge of „epistemic opacity"

Issue raised by Humphreys:[1] **Justification of φ** appears (to some) to be **opaque**.[1,2]

preparatory stage

Newman:[3] prove cor-rectness in advance

e.g., by formal software verification[3]

(usually inapplicable)

validation becomes „holistic"

alleged sources of opacity

no explicit justification is computed

stochastic elements (Monte Carlo)

Lyapunov instability (molecular dynamics)

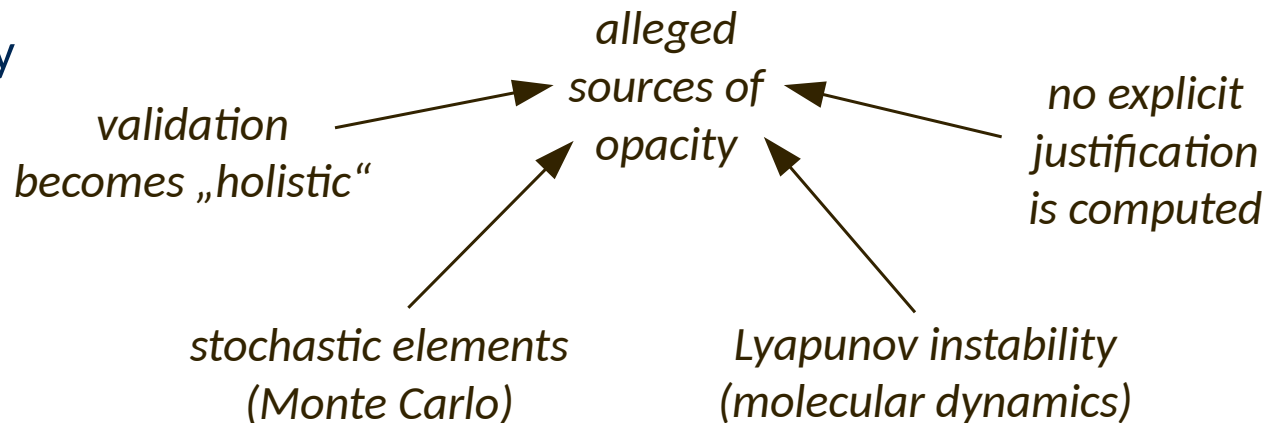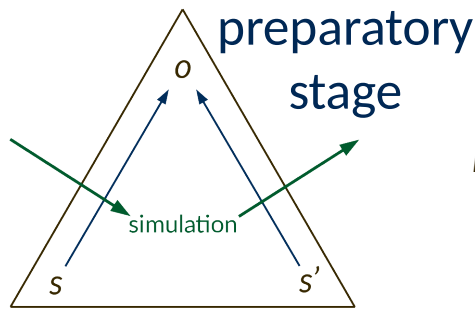However, experiments are not usually regarded as „opaque."

[1]Knowledge of „*all epistemically relevant elements*" cannot be attained (Humphreys, 2004, 2011).
[2]Durán and Formanek (2018): „*epistemically relevant elements*" = „*steps of the [...] justification*".
[3]Required for non-opacity by Newman (2016), a requirement criticized by Durán & Formanek (2018).

Epistemic grounding: Trust by understanding provenance

Justification by epistemic grounding:[1]

preparatory stage

ex ante
Planning

execution stage

in actu
Steering

„Reflexion im Vollzug"[2]

Durán and Formanek:[1] „Computational reliabilism"

[1]J. M. Durán, M. Formanek, *Minds and Machines* 28(4), 645–666, **2018**.
[2]Tulatz, *Epistemologie als Reflexion wissenschaftlicher Praxen*, **2018**.

Epistemic FAIRness by provision of provenance metadata

Three modes of justification by epistemic grounding:

preparatory
stage

o

simulation

s   *s'*

ex ante
Planning

execution
stage

o

simulation

s   *s'*

in actu
Steering

„Reflexion im Vollzug"[1]

critical
analysis

o

simulation

s   *s'*

ex post
Evaluation

„Reflexion des Vollzugs"[1]

Epistemic opacity is reduced by **epistemic FAIRness**, *i.e.*, the FAIR provision of a provenance description via a research data infrastructure that permits a reevaluation of the research workflow over an open epistemic space.

[1]Tulatz, *Epistemologie als Reflexion wissenschaftlicher Praxen*, **2018**.

MODA-based provenance metadata for simulation results[1]

electronic
- EL.1: Ab-initio quantum mechanics
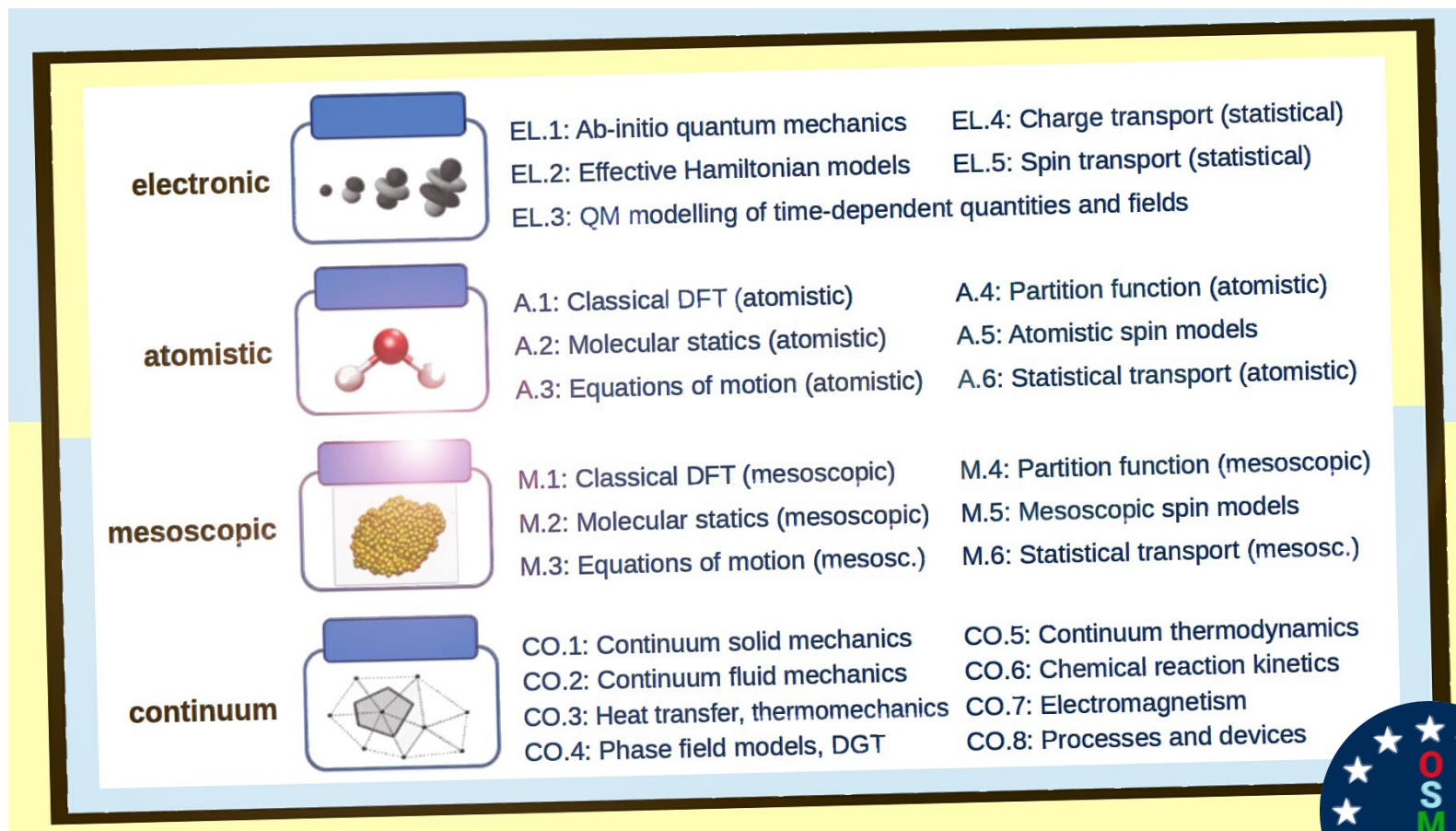- EL.2: Effective Hamiltonian models
- EL.3: QM modelling of time-dependent quantities and fields
- EL.4: Charge transport (statistical)
- EL.5: Spin transport (statistical)

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

MODA-based provenance metadata for simulation results

simulation workflow W

classical mechanical equations of motion (atomistic-mesoscopic)

M_1

materials relation: LJTS potential

continuum thermodynamics

M_3

continuum fluid mechanics and heat transfer

M_2

starting point of W

continuum thermody...

M

MD simulation with ls1 mardyn

starting point of C_2

HEM solver

starting point of C_1

DGSEM with FLEXI

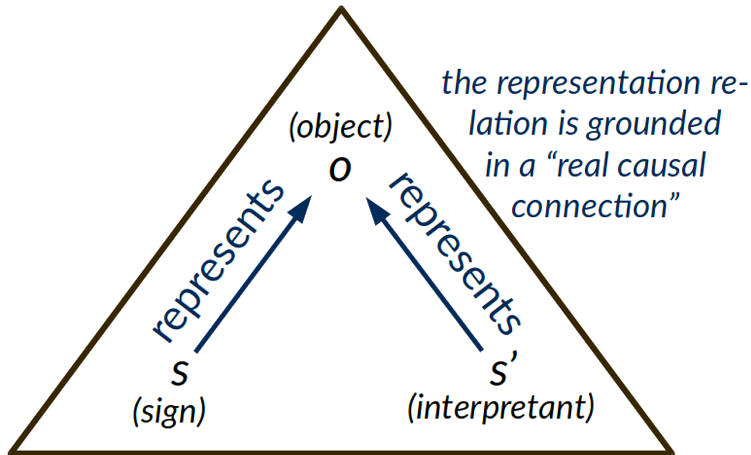use case: Riemann problem

simulation outcome

OSMO-based **provenance description** as an extension of the MODA workflow meta-data standard:

For all elements of the graph notation, there are corresponding concepts and relations from the ontology OSMO.
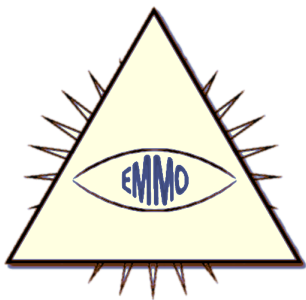
Ontology for Simulation, Modelling, and Optimization (OSMO)
J. Chem. Eng. Data 65(3), 1313–1329, 2020

EMMO-based provenance metadata for simulation results

Peircean semiotics

(object)
o

represents   represents

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

s
(sign)

s'
(interpretant)

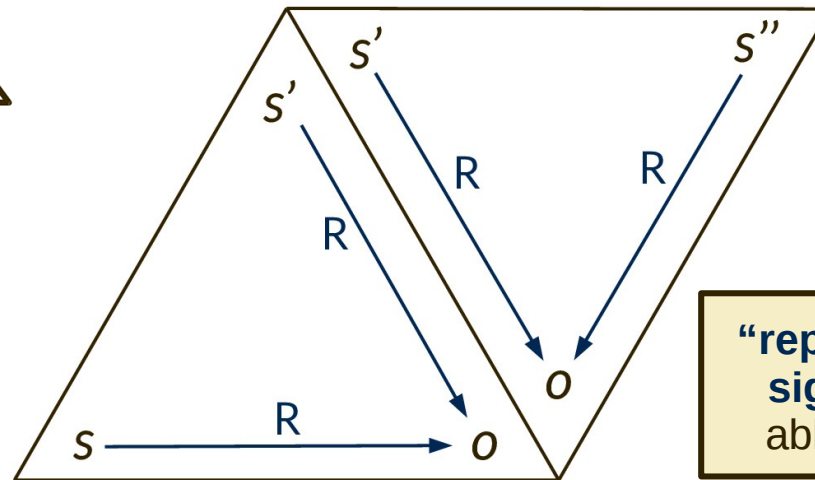*the semiosis, a process by which a new representamen, the interpretant, is created*

EMMO

C. S. Peirce

Cognitive process (example):

- First, experimental data s for the 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).

In Peircean semiotics, a representation relation is carried over from one cognitive step (*i.e.*, triad) to the next.[1]
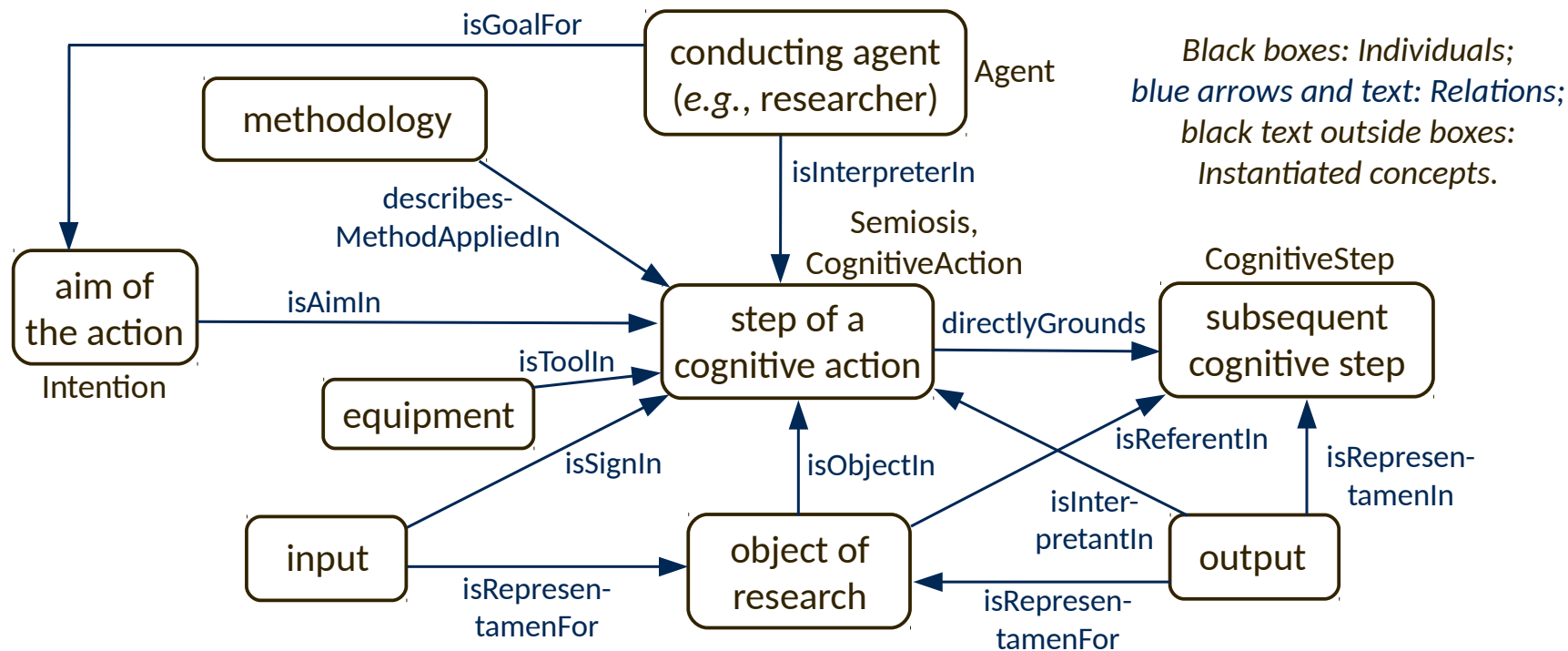
s' s''

s'

R R

R

s R o o

"represents" or **"is sign for"** is here abbreviated by **R**

EMMO-based provenance metadata for simulation results

Schema from the PIMS interoperability infrastructure[1, 2, 3] (PIMS-II)

[1]Introduction and relation to EMMO and MODA discussed by H. A. Preisig *et al.*, DAMDID, submitted, **2021**.
[2]PIMS-II OWL ontology for cognitive processes accessible at http://www.molmod.info/semantics/pims-ii.ttl.
[3]Modal first-order logic ontology at 10.5281/zenodo.4849611; examples at doi:10.5281/zenodo.4679522.

EMMO-based provenance metadata for simulation results

Mereosemiotics:[1, 2, 3] Combination of mereotopology and Peircean semiotics

[1]M. T. Horsch, S. Chiacchiera, B. Schembera, M. A. Seaton, I. T. Todorov, *Proc. WCCM-ECCOMAS 2020*, **2021**.
[2]First-order logic implementation, doi:10.5281/zenodo.4849611; examples, doi:10.5281/zenodo.4679522.
[3]Introduction and relation to EMMO and MODA discussed by H. A. Preisig *et al.*, DAMDID 2021, submitted.

Knowledge representation
for materials modelling marketplaces

M. Horsch,[1, 2] S. Chiacchiera,[2] M. Seaton,[2] I. Todorov,[2] P. Klein,[3] N. Konchakova,[4] C. Niethammer,[1] J. Vrabec,[5] D. Toti,[6, 7] G. Mogni,[7] G. Goldbeck,[7] P. Schiffels,[8] and W. Cavalcanti,[8] *High Performance Computing Center Stuttgart,[1] UKRI STFC Daresbury Laboratory,[2] Fraunhofer ITWM,[3] Helmholtz-Zentrum Hereon,[4] Techn. Univ. Berlin,[5] Catholic Univ. Brescia,[6] Goldbeck Consulting Ltd.,[7] Fraunhofer IFAM[8]*