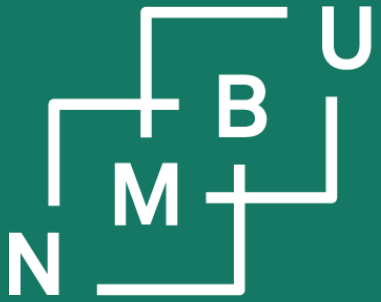
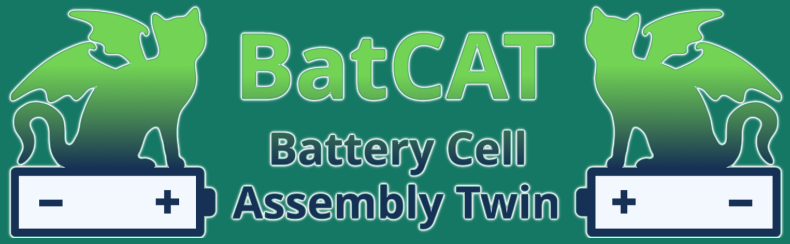
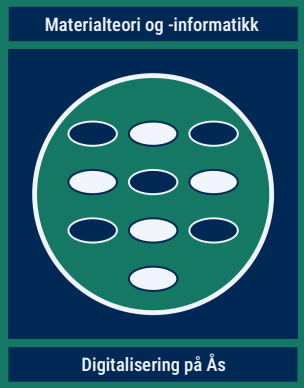






Scope of physics-based simulation artefacts

Martin T. Horsch,[1] Fadi Al Machot,[1] and Jadran Vrabec[2]

DCLXVI 2024 Workshop, 11.12.2024, Kaiserslautern

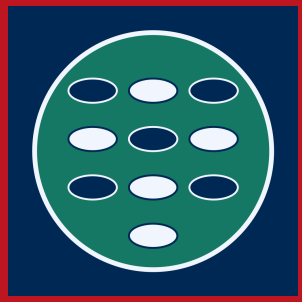
[1]Material Theory and Informatics Group, Norwegian University of Life Sciences, Ås
[2]Thermodynamics Group, Technische Universität Berlin

Fakultet for realfag og teknologi

Forskergruppe materialteori og -informatikk

Noregs miljø- og biovitskaplege universitet

Materialteori og -informatikk



Digitalisering på Ås

BatCAT Battery Cell Assembly Twin

DigiPass Know your products!

KGA

1. Epistemic metadata
2. *Two case studies*
3. *Simulation scope*
4. *Simulation artefacts*
5. *Mid-level ontologies*

Opacity vs. transparency

European AI Act: "To address concerns related to **opacity** and […] fulfil their obligations under this Regulation, **transparency** should be required for high-risk AI systems before they are placed on the market […]. High-risk AI systems should […] enable deployers to understand how the AI system works […]. High-risk AI systems should be accompanied by **appropriate information**".

Epistemic opacity can occur when simulation-based and data-driven methods are used. The concept was introduced by **Humphreys** in *Extending Ourselves*[1] (2004), developed further in later work,[2] and has had a substantial impact.[3]

> **Epistemic opacity** (Humphreys, 2011): A «process is **epistemically opaque** relative to a cognitive agent *X* at time *t* […if …] *X* does not know at *t* all of the **epistemically relevant elements**»[2]

[1]P. Humphreys, *Extending Ourselves Computational Science, Empiricism, and Scientific Method*, **2004**.

[2]P. Humphreys, in M. Carrier, A. Nordmann, *Science in the Context of Application*, *pp.* 131–142, Springer, **2011**.

[3]J. M. Durán, N. Formanek, *Minds and Machines* **28**(4): 645–666, doi:10.1007/s11023-018-9481-6, **2018**.

Epistemic metadata

Epistemic metadata are the information that **establishes the knowledge status** of data or digital objects.[1]

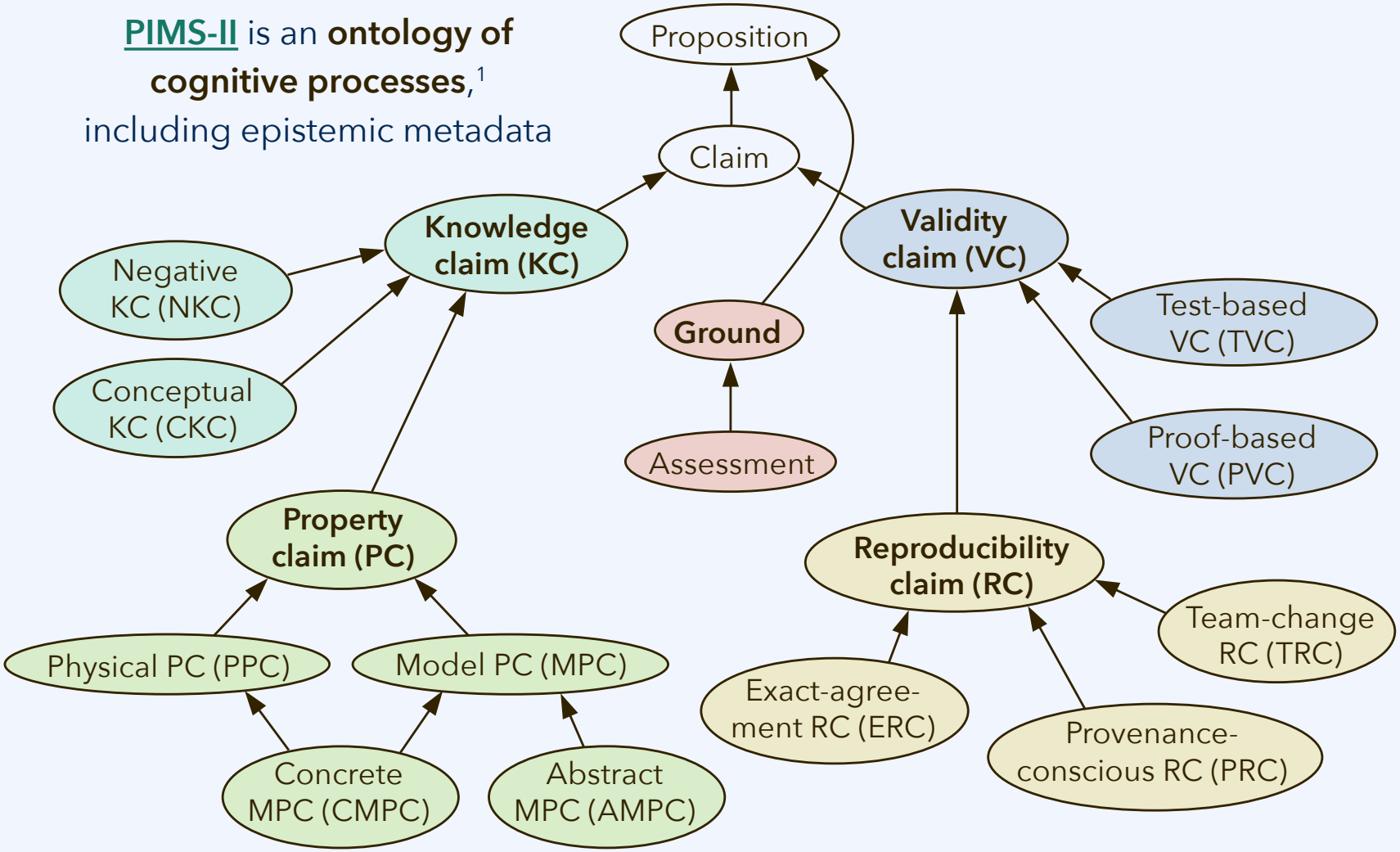
> **Questions we must answer to establish the knowledge status:**
>
> **a)** "what **knowledge claim** φ has been formulated?,"
> **b)** "where do the data and the claim come from?" (**provenance**),
> **c)** "what **validity claim** was made about φ?,"
> **d)** "why should we accept any of this?" (**grounding**).

Key epistemic metadata items are the **knowledge claims** made based on data, their **provenance**, **validation** and **reproducibility**, and **epistemic grounding**.

In *Proc. JOWO 2022*, CEUR *vol.* **3249**: *p. 2 (CAOS)*, CEUR-WS, **2022**.
In *Proc. ICAPAI 2023*, doi:10.1109/icapai58366.2023.10193944, IEEE, **2023**.
In *Proc. FOIS 2023*, *pp.* 302–319, doi:10.3233/faia231136, IOS, **2023**.

Mid-level ontology of epistemic metadata

PIMS-II is an **ontology of cognitive processes**,[1] including epistemic metadata



[1]OWL implementation under http://www.molmod.info/semantics/pims-ii.ttl

Refactoring of mid-level ontology

PURL for the new system, MSO-EM (ontologies for **modelling, simulation, optimization**, and **epistemic metadata**), which is under construction:

https://www.purl.org/mso-em

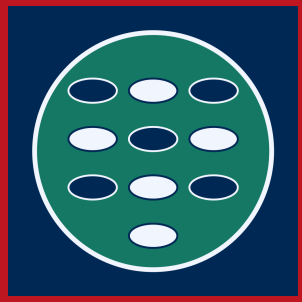
BatCAT organizational github: https://github.com/HE-BatCAT

Design principles:

- Strong alignment with DOLCE (through DOLCE Lite)
- OWL2 EL profile expressivity level
- Ongoing development, with easy stable access to versioned releases
- Simple modules, each with maximum three taxonomy levels and maximum three top concepts
- Backwards compatibility with equivalences to the preceding mid-level ontology development (PIMS-II) to the maximum possible extent
- All modules of the ontology are directly aligned with DOLCE

Noregs miljø- og biovitskaplege universitet

Materialteori og -informatikk



Digitalisering på Ås

1. Epistemic metadata
2. Two case studies
3. *Simulation scope*
4. *Simulation artefacts*
5. *Mid-level ontologies*

Case study on knowledge claims

Epistemic metadata and their documentation were explored for the domain of molecular modelling and simulation within engineering thermodynamics:

First stage report (10 cases), doi:10.5281/zenodo.7516532, **2023**.

Discussion of *five papers each* from *two research groups* (London, Berlin) without involving the papers' authors. Obtained a tentative **taxonomy for epistemic metadata**, later implemented into the PIMS-II ontology.

Second stage report (12 claims), doi:10.5281/zenodo.7608074, **2023**.

Discussion of *two claims each* from *six papers*, with two papers each from three research groups (London, Berlin, Kaiserslautern), involving the papers' authors. Discussed aspects such as the **grounding of knowledge claims** with authors.

Case study on knowledge claims

Grounding of claims: See also the reference ontology of trust **ONTrust**[1, 2]

	trust	reliance
Type-1 The **results** establish their own validity.	*Typical:* Mathematical argument (proof) over of a conceptual framework designed around widely accepted definitions and axioms.	*Schema:* A new theory is more reliable because it is simpler, covers more phenomena, or represents underlying physics. (**theoretical virtues**)
Type-2 The **provenance** of the results tells that they are valid.	*Case study example:* Chatwell and Vrabec argue: It is OK to use a cutoff radius of 5.5σ for the LJ potential, since this was done in three cited works from the literature.	*Typical:* We used a model, method, and simulation code validated in the past and – usually – very accurate. (**process reliabilism**)

[1]Baratella *et al.*, «The many facets of trust», in *Proc. FOIS 2023*, doi:10.3233/faia231115, **2024**.

[2]https://github.com/unibz-core/trust-ontology

Case study on objects and objectives*

abstractness of object

qualitative reference model

really existing, simulations and experiments

Homes *et al.*

liquid slab evaporation

Abstract object: The LJTS fluid.
Abstract objective: Understanding evaporation phenomena.

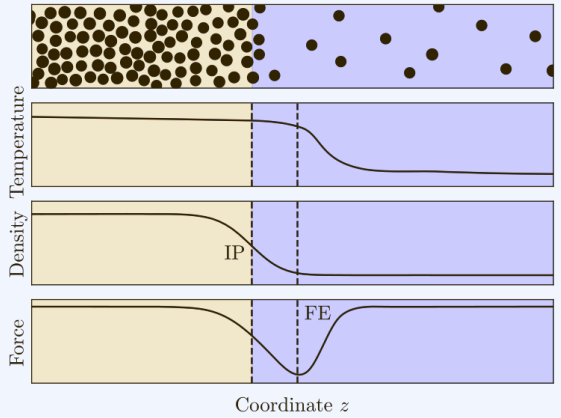


Figure 1. Cut through the simulation volume with profiles of temperature, density and force over the spatial coordinate *z*. Background colours mark different fluid regions

Table 8. Comparison of the Molar Volume Calculated Using eq 4 with Experimental Literature Data and Molecular Simulation Results for the Ternary and Quaternary Mixtures at 298.15 K and 0.101 MPa

mixture	no. data points	AARD %	source
water + methanol + ethanol	14	0.24	experiment[11]
	8	0.65	experiment[18]
	36	0.25	experiment[10]
	13	1.20	experiment[19]
	5	0.80	experiment[20]
	3	0.29	simulation, this work
	27	0.43	simulation[4]

Concrete object and concrete objective

Guevara *et al.*

water-alcohol mixtures

abstractness of objective

parameter variation, prediction, comparison to experiment

qualitative understanding and characterization of an idealized model

*Sample: Nine papers published by the TU Berlin Thermodynamics group since 2021 with the most citations.

Case study on objects and objectives

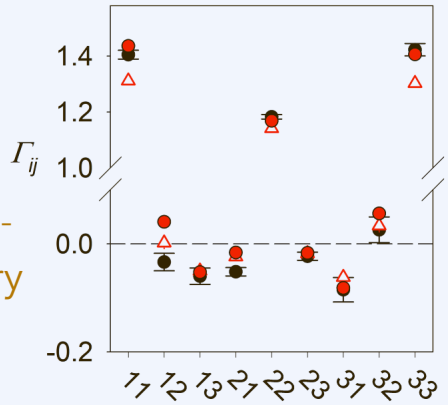
abstractness of object

qualitative reference model

varying boundary conditions, all plausibly really existing

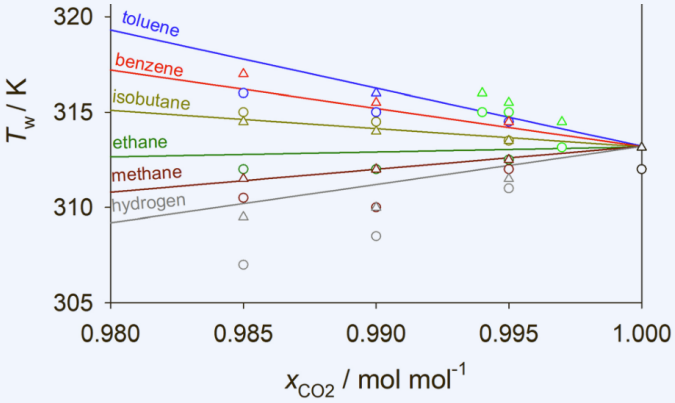
really existing, simulations and experiments

Fingerhut *et al.*
Fig. 3
quaternary diffusion



Homes *et al.*
liquid slab evaporation

Liquid-like super-critical quaternary LJ mixture



Widom line of binary mixtures with supercritical CO_2 as main component

Šarić *et al.*
scCO$_2$ solutions

Fig. 8 Mole fraction dependence of the Widom line temperature along the isobar p = 9 MPa employing two criteria

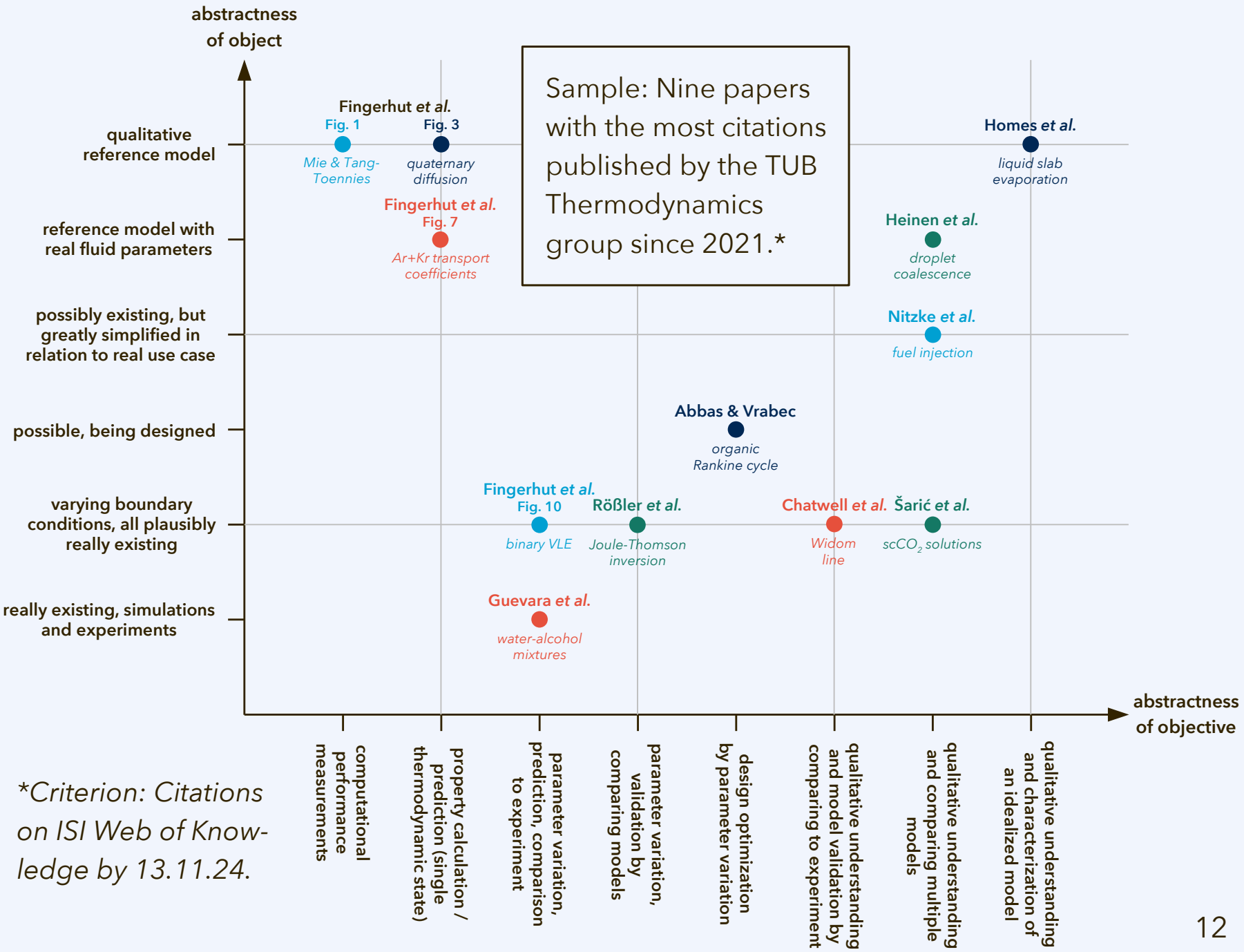
abstractness of objective

property calculation / prediction (single thermodynamic state)

parameter variation, prediction, comparison to experiment

qualitative understanding and comparison of multiple models

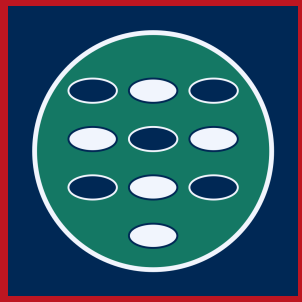
qualitative understanding and characterization of an idealized model



**Criterion: Citations on ISI Web of Knowledge by 13.11.24.*

Noregs miljø- og biovitskaplege universitet

Materialteori og -informatikk



Digitalisering på Ås

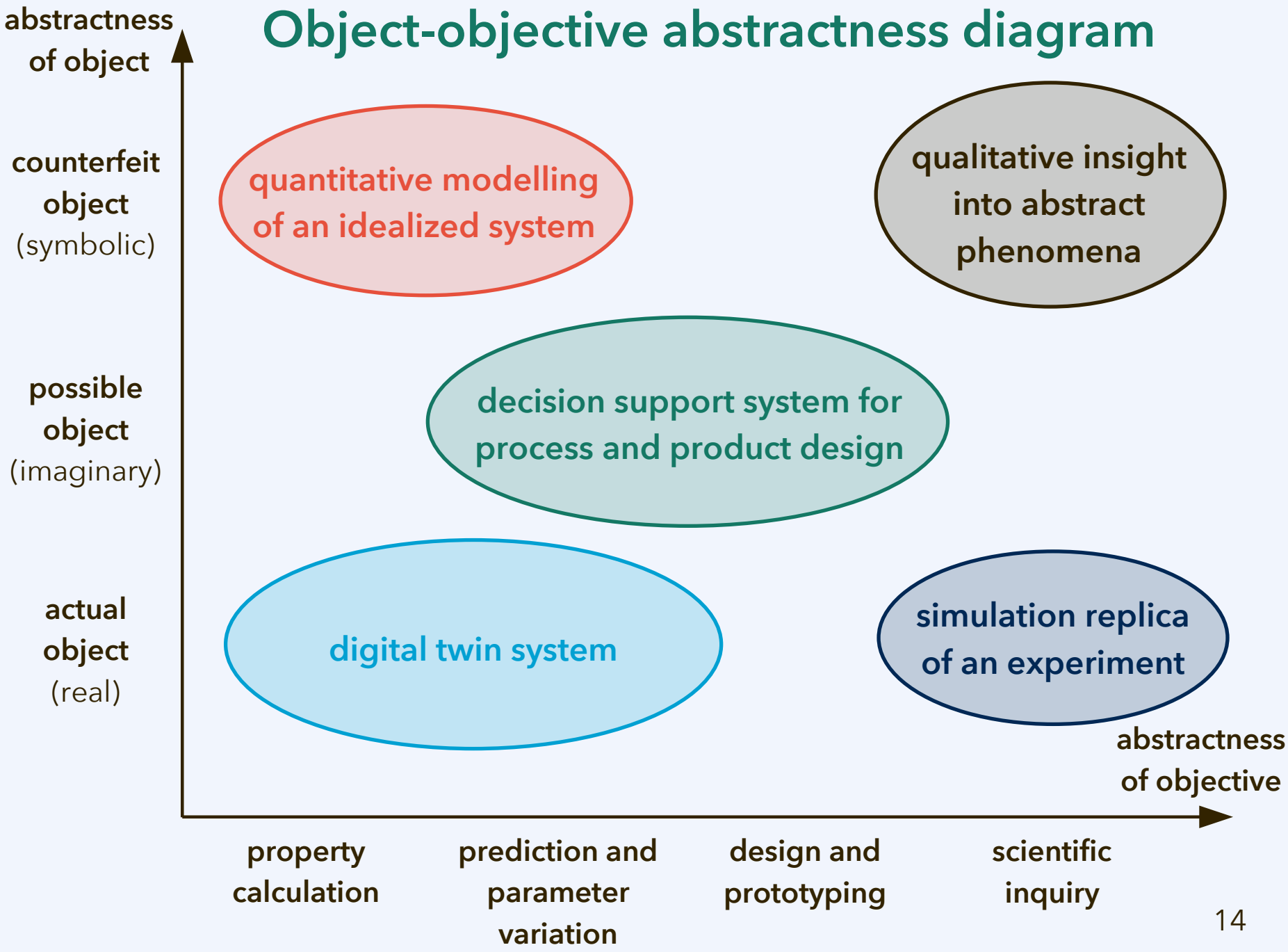
BatCAT Battery Cell Assembly Twin

DigiPass Know your products!

KGA

1. Epistemic metadata
2. Two case studies
3. Simulation scope
4. *Simulation artefacts*
5. *Mid-level ontologies*

Object-objective abstractness diagram



Subject matter as an aspect of scope

Following Yablo,[1] the **subject matter** of a knowledge claim and/or associated research data is given by the **research question that is being answered**, or by the «equivalence relation over logical space» with respect to that question.

Proposition: "*A* is the factually correct answer to question *Q*."
Subject matter of the proposition: *Q*.

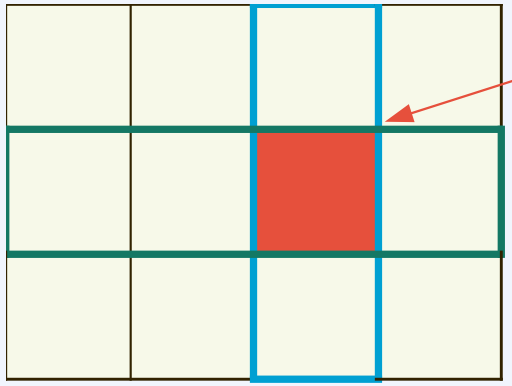
We distinguish two ways of **combining topics**. Related topics q_1 and q_2 form a **topical product** q_1q_2 where the partitioning of logical space by $\equiv_{q1q2}$ is the *product of the sets of equivalence classes* with respect to $\equiv_{q1}$ and $\equiv_{q2}$.

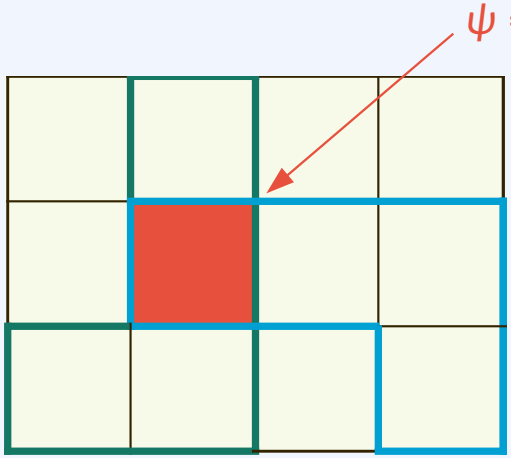
However, long papers *etc.* can also be about many topics that are not closely related. They stand side by side. We call this a **topical sum** $q_1 + q_2$, *e.g.*,

q_1 = a theoretical research question from statistical mechanics,
q_2 = topic of a concrete series of simulations from the same paper.

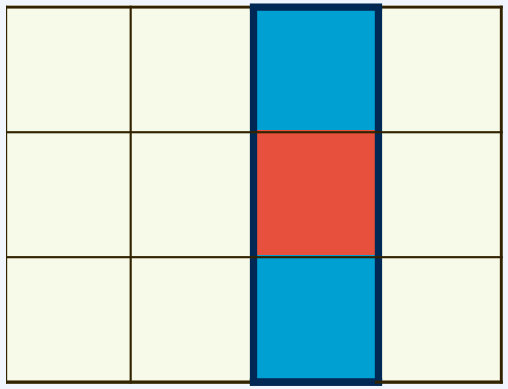
[1]S. Yablo, *Aboutness*, Princeton Univ. Press (ISBN 978-0-691-14495-5), **2014**.

Subject matter as an ingredient of semantics



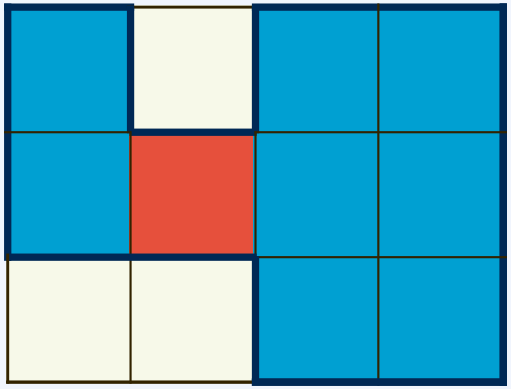


$\psi - \varphi_y = (\varphi_x \wedge \varphi_y) - \varphi_y \equiv \varphi_x$



φ_x is recovered by subtracting φ_y

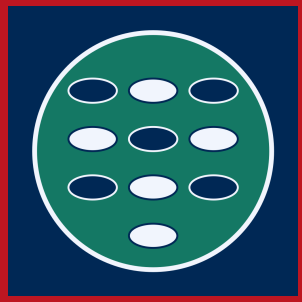
$\psi - \varphi_B = (\varphi_A \wedge \varphi_B) - \varphi_B \not\equiv \varphi_A$



φ_A cannot be perfectly recovered

Noregs miljø- og biovitskaplege universitet

Materialteori og -informatikk

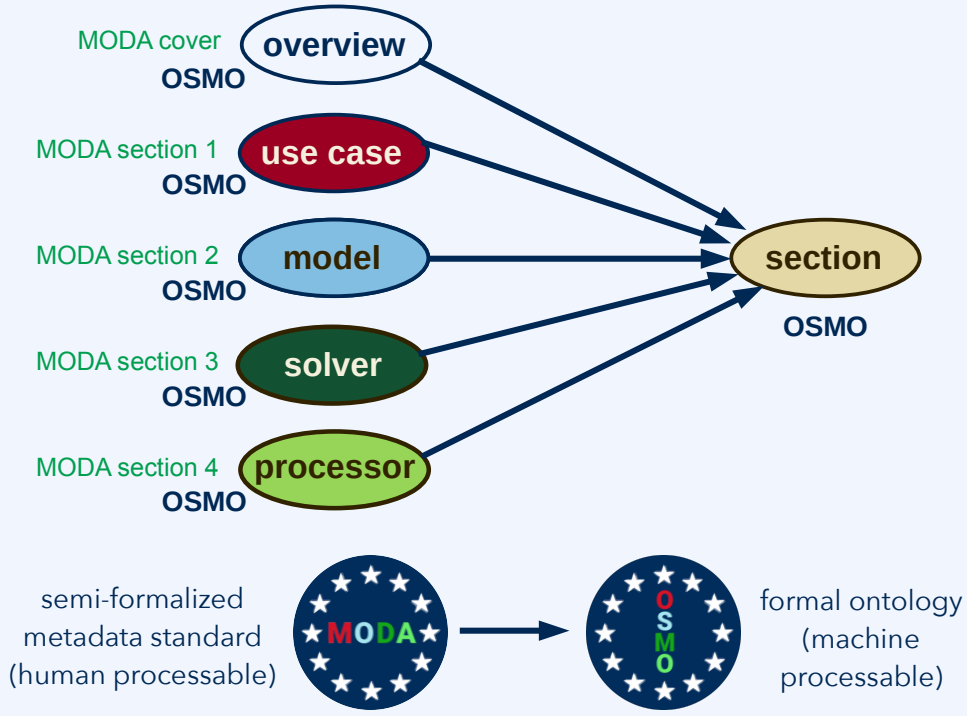


Digitalisering på Ås

1. Epistemic metadata
2. Two case studies
3. Simulation scope
4. Simulation artefacts
5. *Mid-level ontologies*

CWA 17960 MODA

MODA: Well-known metadata standard developed by the EMMC community. We compared its treatment of **simulation artefacts** to that from the CWA ModGra.



CWA 17960 ModGra

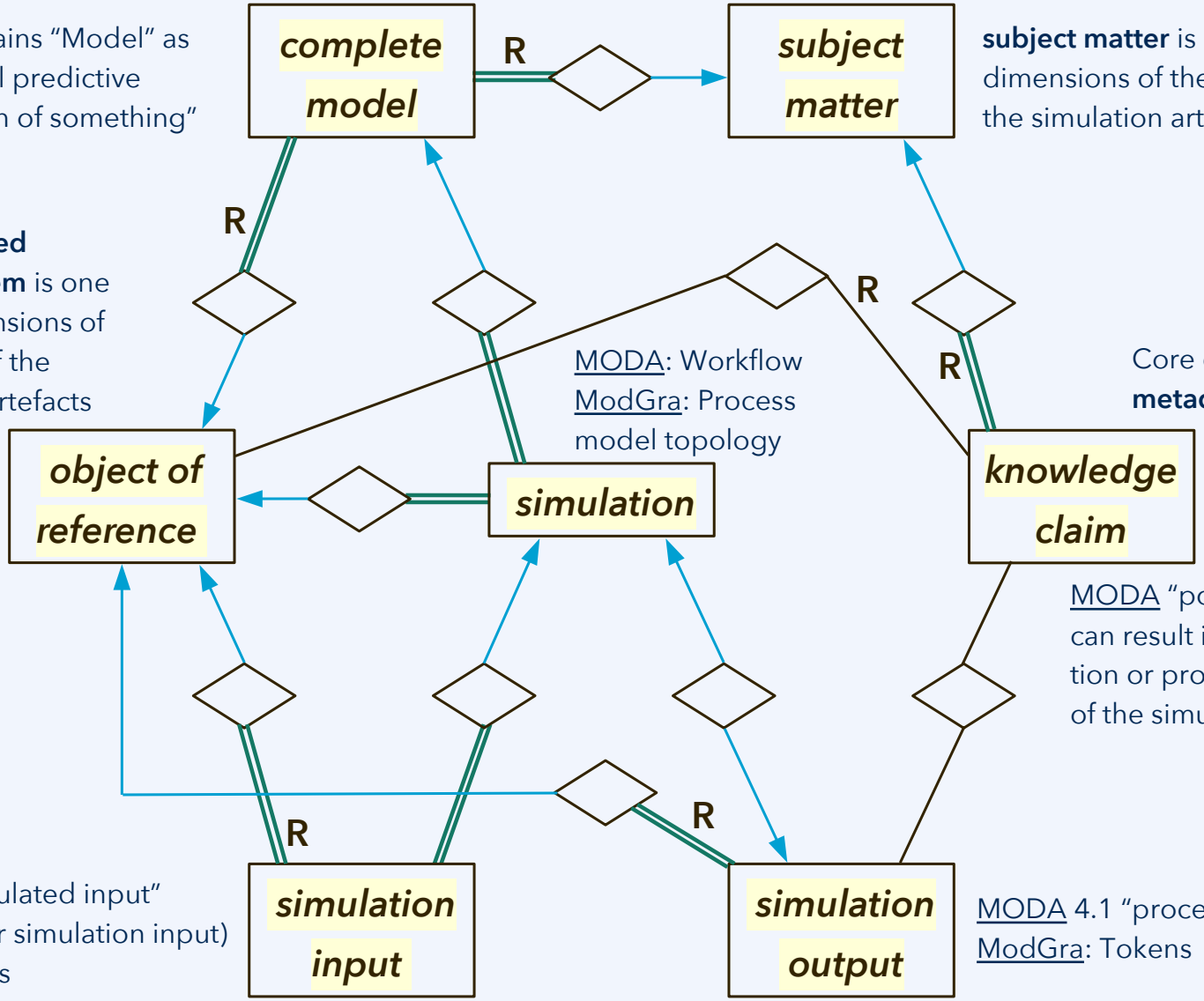
The **process model topology** is a generalized Petri net. Tokens represent **extensive physical quantities** or, alternatively, **data items** that can be exchanged.

The process model's system of equations is represented by the Petri net's **transitions** and the **places** (capacities) which store the balanced quantities, or alternatively the data items.

An **EMMO export**, using **TriG format**, has been developed.

MODA: "Physics-based model" defined as "solvable set of [...] physic[al] equation and one or more materials relations"

ModGra contains "Model" as "mathematical predictive representation of something"



subject matter is one of the dimensions of the **scope** of the simulation artefacts

The **simulated object/system** is one of the dimensions of the **scope** of the simulation artefacts

MODA: Workflow
ModGra: Process model topology

Core **epistemic metadata** item

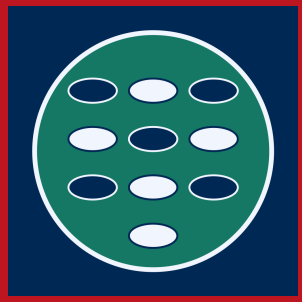
MODA "postprocessing" can result in an interpretation or processed version of the simulation output

MODA 2.5 "simulated input" (may be typo for simulation input)
ModGra: Tokens

MODA 4.1 "processed output"
ModGra: Tokens

Noregs miljø- og biovitskaplege universitet

Materialteori og -informatikk



Digitalisering på Ås

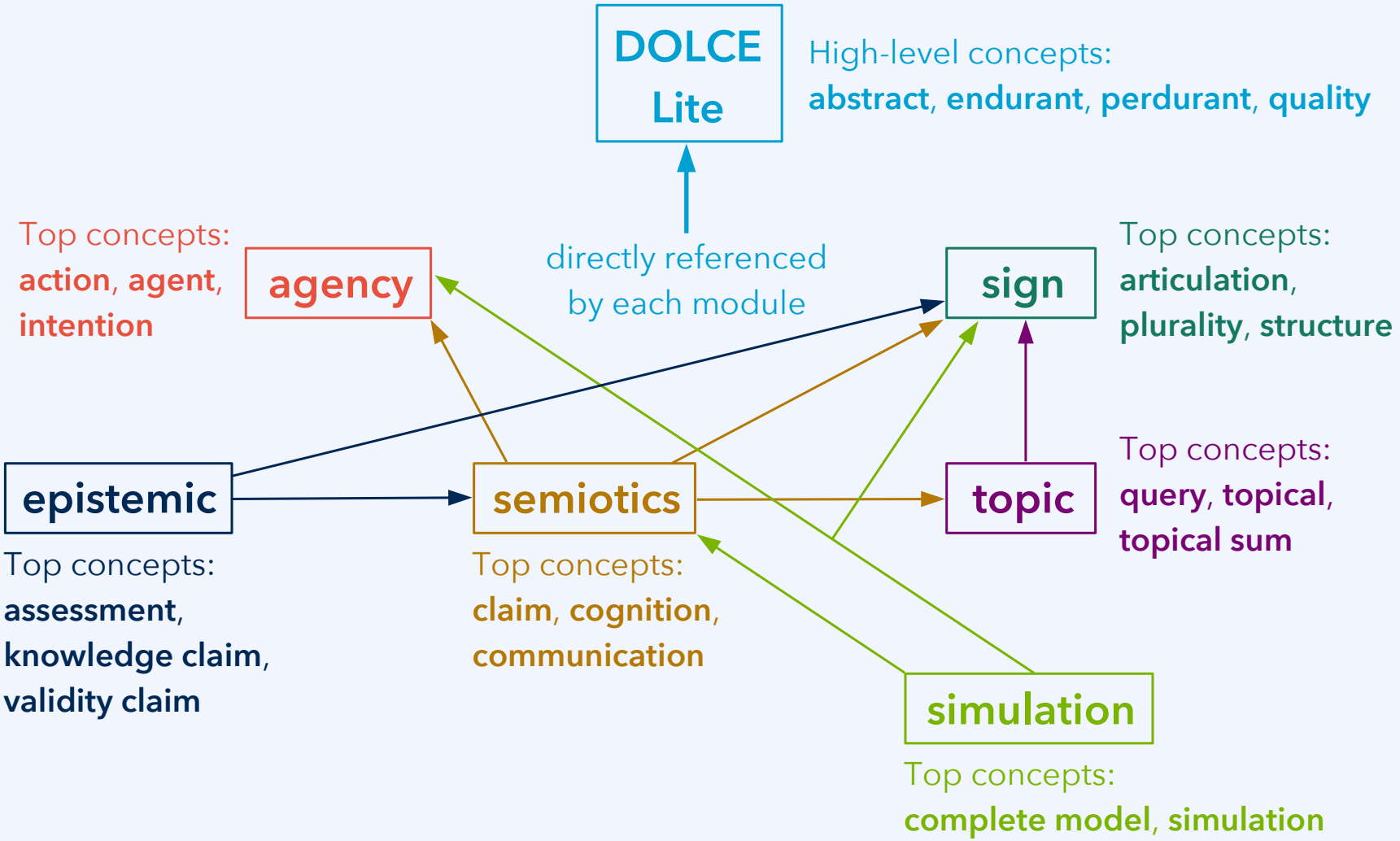
BatCAT Battery Cell Assembly Twin

DigiPass Know your products!

KGA

1. Epistemic metadata
2. Two case studies
3. Simulation scope
4. Simulation artefacts
5. Mid-level ontologies

MSO-EM ontology modules



Relations from the E-R diagram

(1) MSO-EM simulation: **evaluates model**
DOLCE: specific constant dependent

(2) MSO-EM semiotics: **involves sign**
DOLCE: participant

(3) MSO-EM semiotics: **involves referent**
DOLCE: weak connection

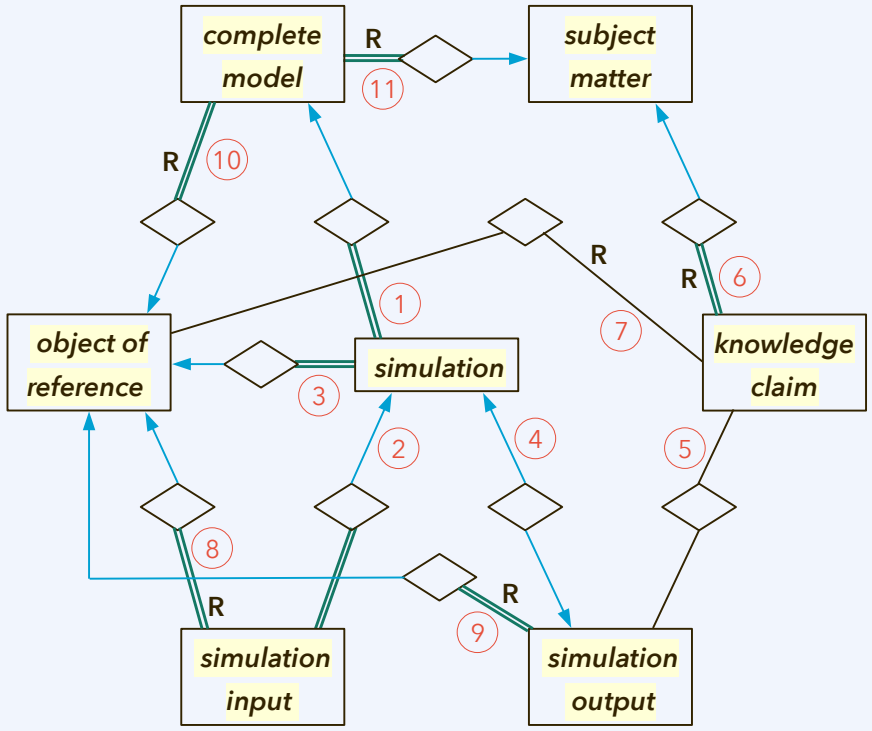
(4) MSO-EM semiotics: **involves interpretant**
DOLCE: participant

(5) MSO-EM semiotics: **is based on**
DOLCE: generic constituent

(6) MSO-EM topic: **has subject matter**
DOLCE: generically dependent on

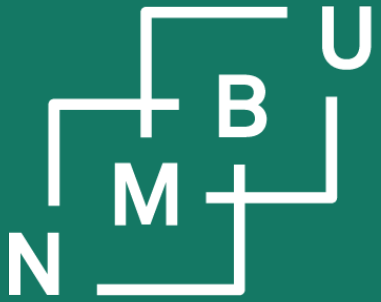
(7) MSO-EM topic: **is about**
DOLCE: generically dependent on

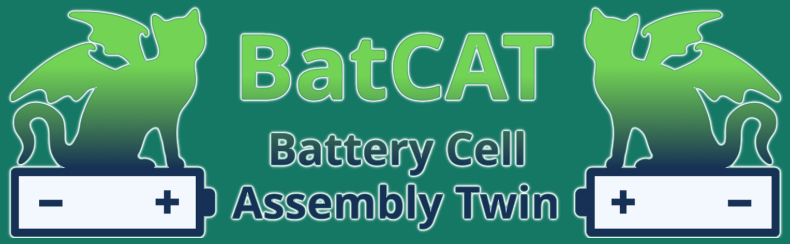
(8, 9) MSO-EM sign: **represents**
DOLCE: generically dependent on

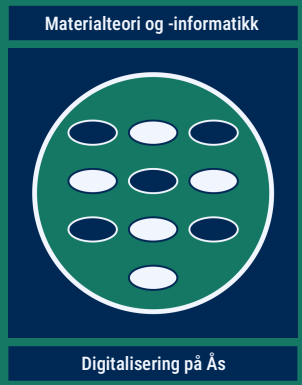


(10) MSO-EM simulation: **articulatesModelOf**
DOLCE: generically dependent on

(11) MSO-EM topic: **is about**
DOLCE: generically dependent on





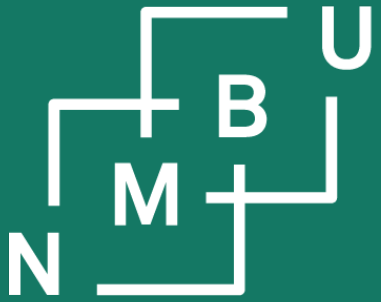


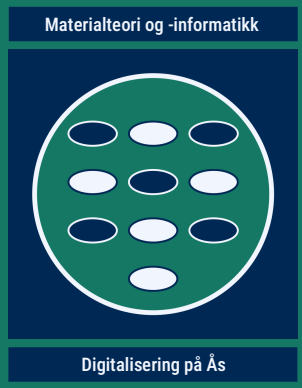
BatCAT has received funding from the European Union's **Horizon Europe** research and innovation programme under **grant agreement no. 101137725**. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the project, the European Climate, Infrastructure and Environment Executive Agency (CINEA), or the European Union. Neither BatCAT nor the CINEA or the EU can be held responsible for them.

Finansiert av Den europeiske union

11th December 2024

NMBU Fakultet for realfag og teknologi Forskergruppe materialteori og -informatikk



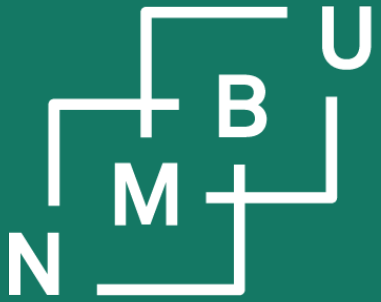


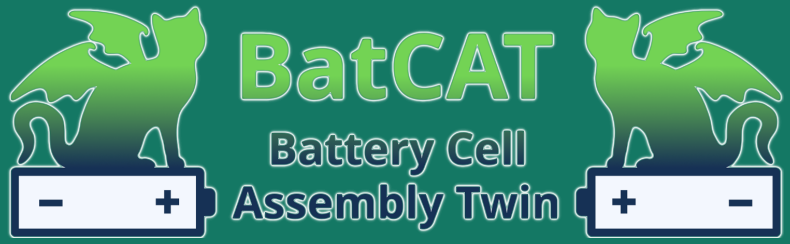
DigiPass CSA has received funding from the European Union's Horizon Europe research and innovation programme under **grant agreement no. 101138510**. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the project, the European Health and Digital Executive Agency (EHDEA), or the European Union. Neither DigiPass CSA nor the EHDEA or the EU can be held responsible for them.

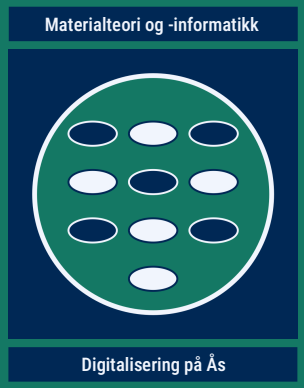
Finansiert av Den europeiske union

11th December 2024

NMBU Fakultet for realfag og teknologi Forskergruppe materialteori og -informatikk







Scope of physics-based simulation artefacts

Martin T. Horsch,[1] Fadi Al Machot,[1] and Jadran Vrabec[2]

DCLXVI 2024 Workshop, 11.12.2024, Kaiserslautern

[1]Material Theory and Informatics Group, Norwegian University of Life Sciences, Ås
[2]Thermodynamics Group, Technische Universität Berlin

Fakultet for realfag og teknologi

Forskergruppe materialteori og -informatikk