Martin Horsch, Silvia Chiacchiera,
Michael Seaton, and Ilian Todorov
STFC Daresbury Laboratory
UK Research and Innovation

Reliable and inter-operable computational molecular engineering

16th July 2019

ICIAM 2019

València

VIMMP
VIRTUAL MATERIALS MARKETPLACE

Science & Technology Facilities Council | UK Research and Innovation

VIMMP VIRTUAL MATERIALS MARKETPLACE

Chemical Engineering Data Science: Session II

Translation and Translators

- Martin Horsch, Silvia Chiacchiera, Michael Seaton, and Ilian Todorov, *“Reliable and interoperable computational molecular engineering”*
- Graziano Frungieri, Gianluca Boccardo, and Marco Vanni, *“Multiscale simulations of industrial problems in an open simulation platform: The compounding of rubber materials as a case study”*
- Patrice Malfreyt, *“Multiscale modelling of polymer materials: Recent advances and challenges”*

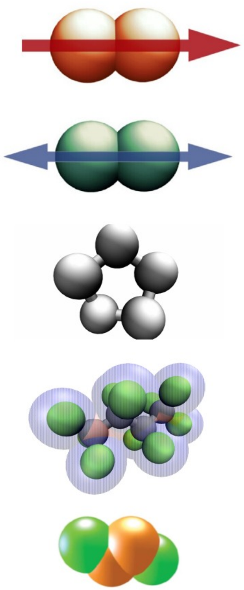
EMMC

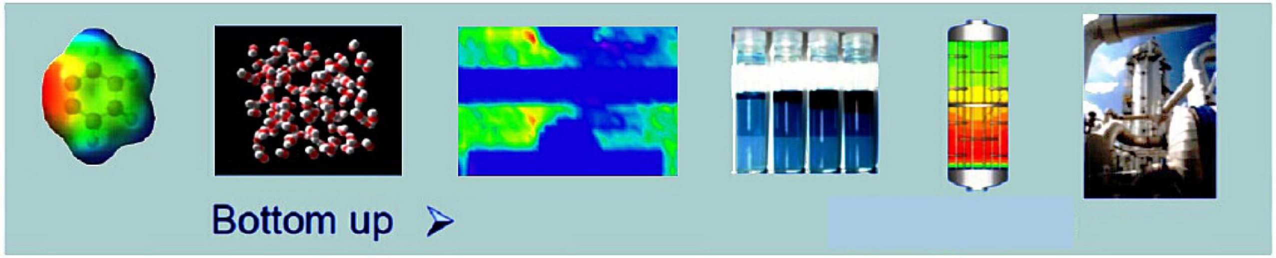
VIMMP
VIRTUAL MATERIALS MARKETPLACE

FORCE MarketPlace

For publication on http://emmc.info/, if desired, send your slides to martin.horsch@stfc.ac.uk

Computational molecular engineering



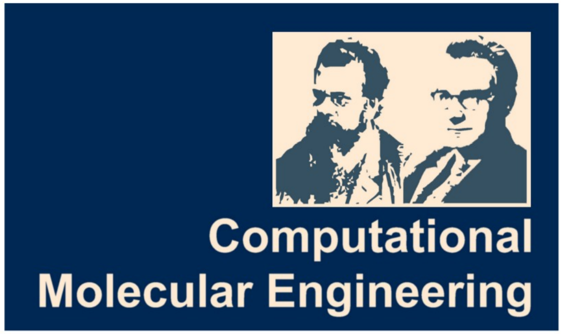


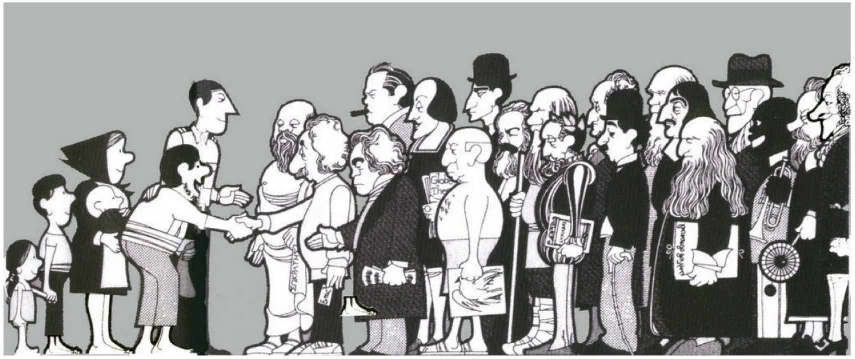
physics-driven aspects
(qualitative validity)

→

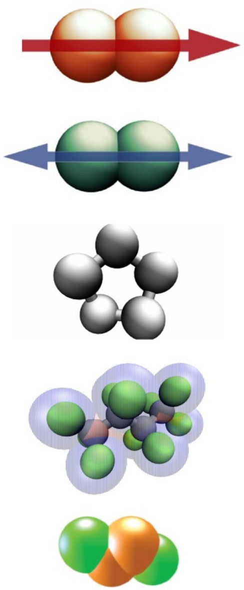
data-driven aspects
(quantitative reliability)

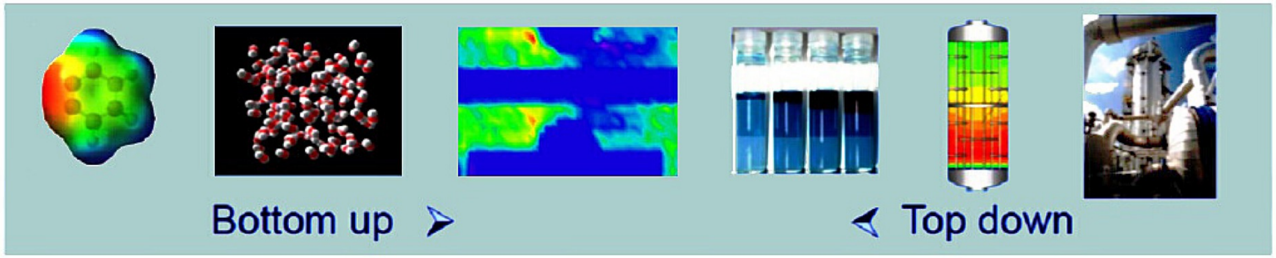
- Realistic representation of underlying physical features





Computational molecular engineering





physics-driven aspects (qualitative validity) ⟷ **data-driven aspects (quantitative reliability)**

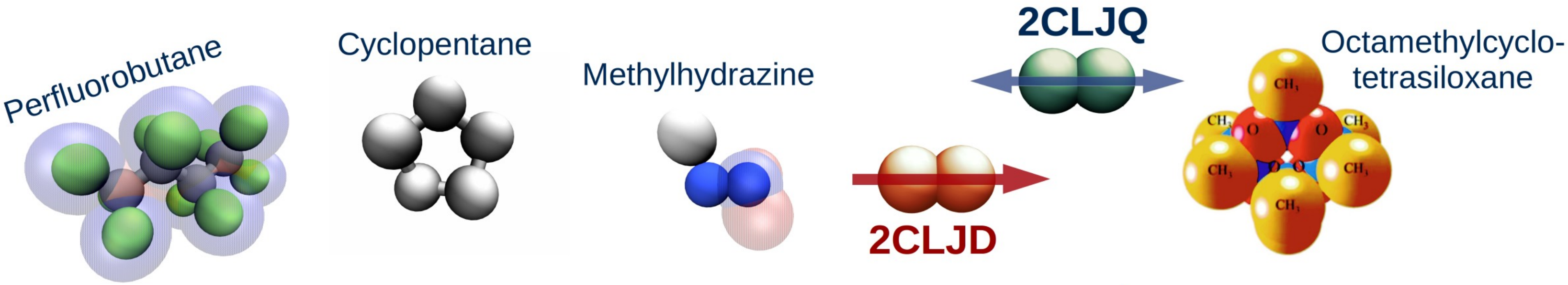
- Realistic representation of underlying physical features

- Models with free parameters that can be adjusted to experimental data
- Reliable interpolation, extrapolation and prediction of thermophysical properties

Computational Molecular Engineering

- **Process Engineering Thermodynamics**
- **Engineering Data Science and Technology**
- **Scientific and High-Performance Computing**

Computational molecular engineering: Model database



Boltzmann-Zuse MolMod DB[1]

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

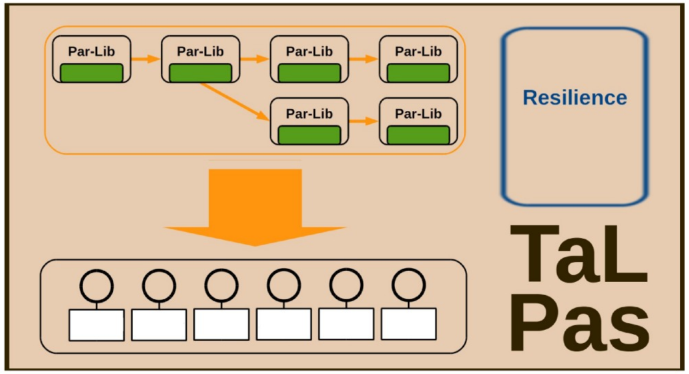
nomenclature and semantic asset development on the basis of the EMMC Review of Materials Modelling (RoMM)

pair potentials for over 150 molecular fluids

Computational Molecular Engineering

Input files can be downloaded for Gromacs, ms2, LAMMPS, and ls1

VIMMP
VIRTUAL MATERIALS MARKETPLACE



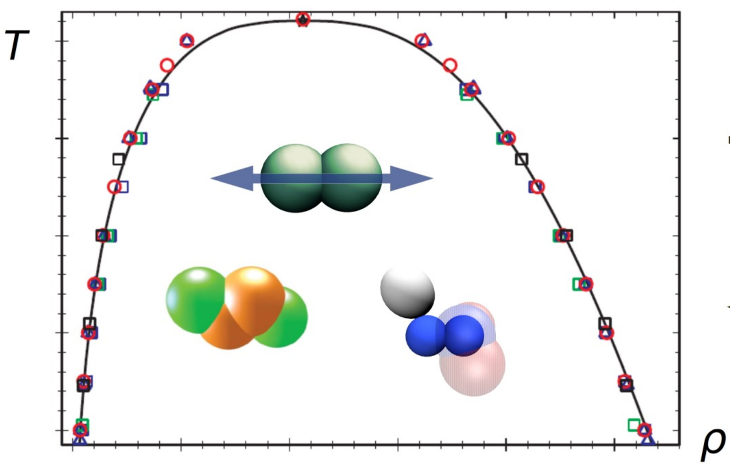
[1]S. Stephan, M. Horsch, J. Vrabec, H. Hasse, *Mol. Sim.* 45(10), 806 – 814, **2019**.

Validation of the molecular models from the MolMod DB

Force field parameterization and validation:

- Pair potentials for fluids are usually adjusted to VLE bulk data, but not to interfacial properties, which can be used for validation.
- Many of the low-molecular fluid models are of the 2CLJD/Q types, with four model parameters: LJ parameter σ, LJ parameter ε, elongation L, and point-multipole strength.

VLE bulk properties

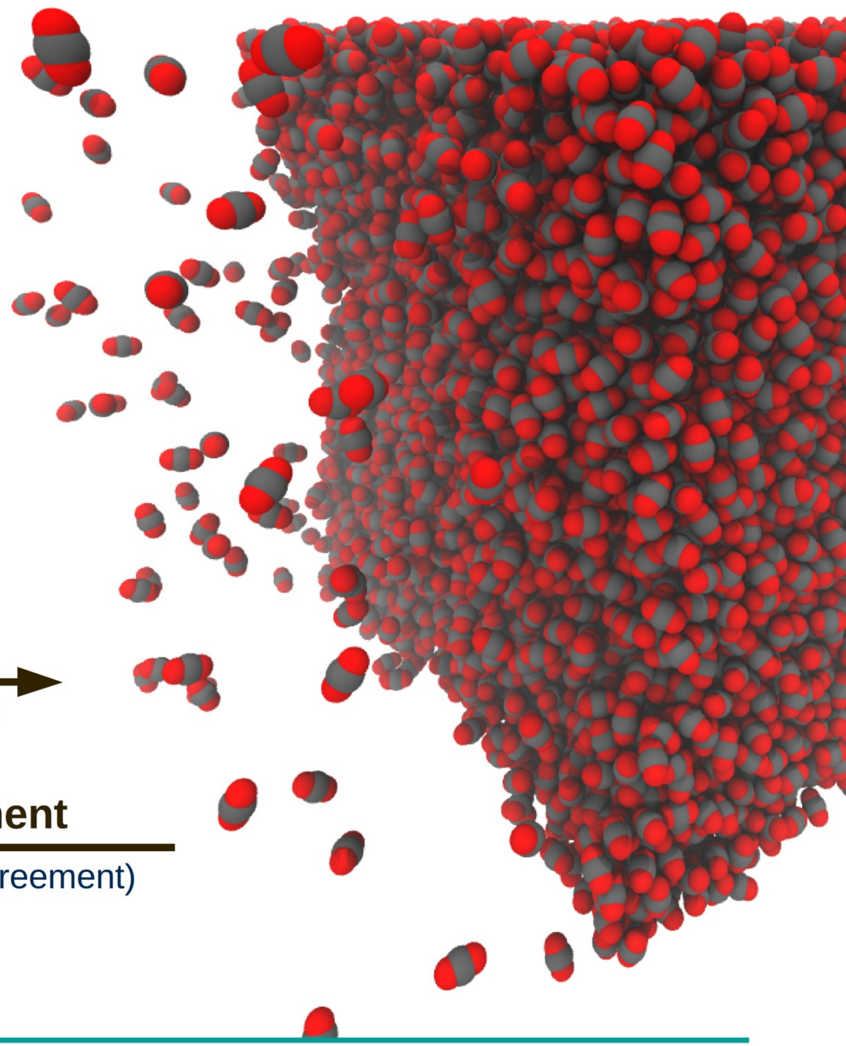


prediction

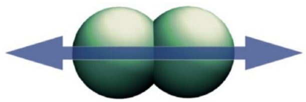
(model property)

assessment

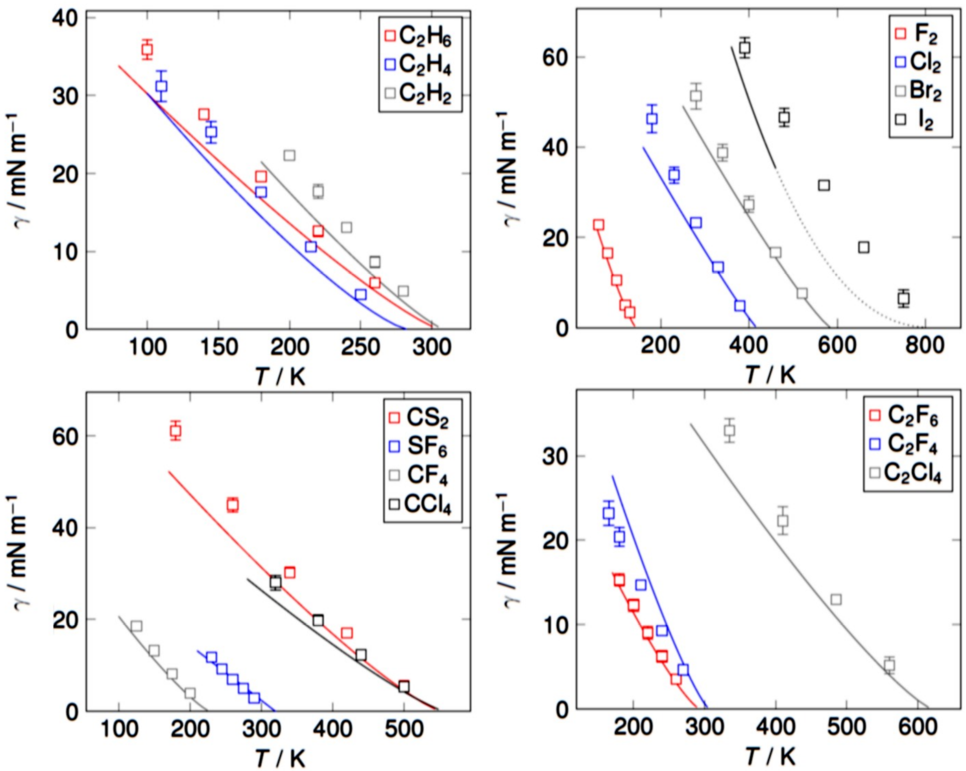
(validation of agreement)

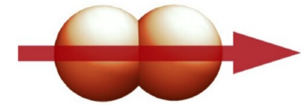


Validation of the molecular models: Surface tension

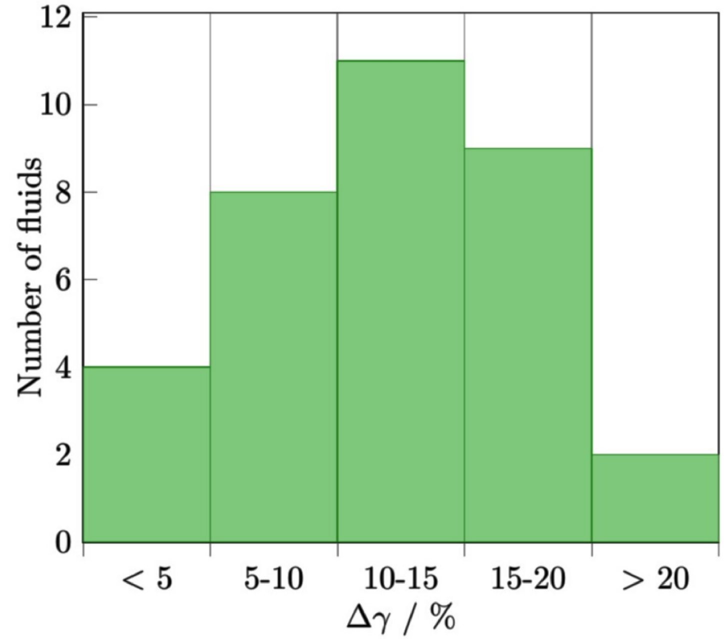


2 LJ Centres + Quadrupole[1]





2 LJ Centres + Dipole[2]

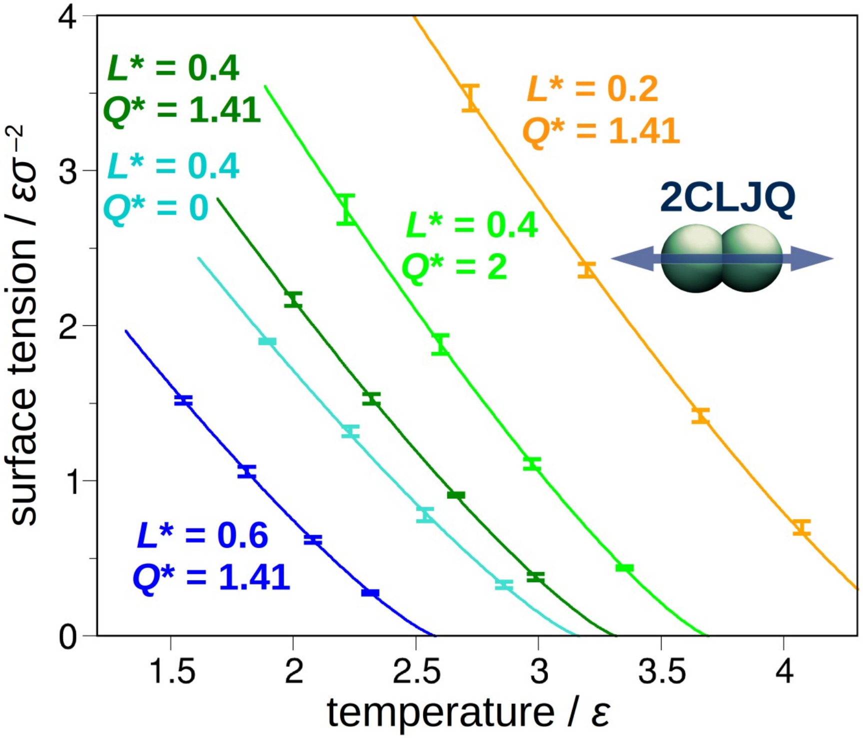


[2]S. Werth, M. Horsch, H. Hasse, *J. Chem. Phys.* 144, 054702, **2016**.

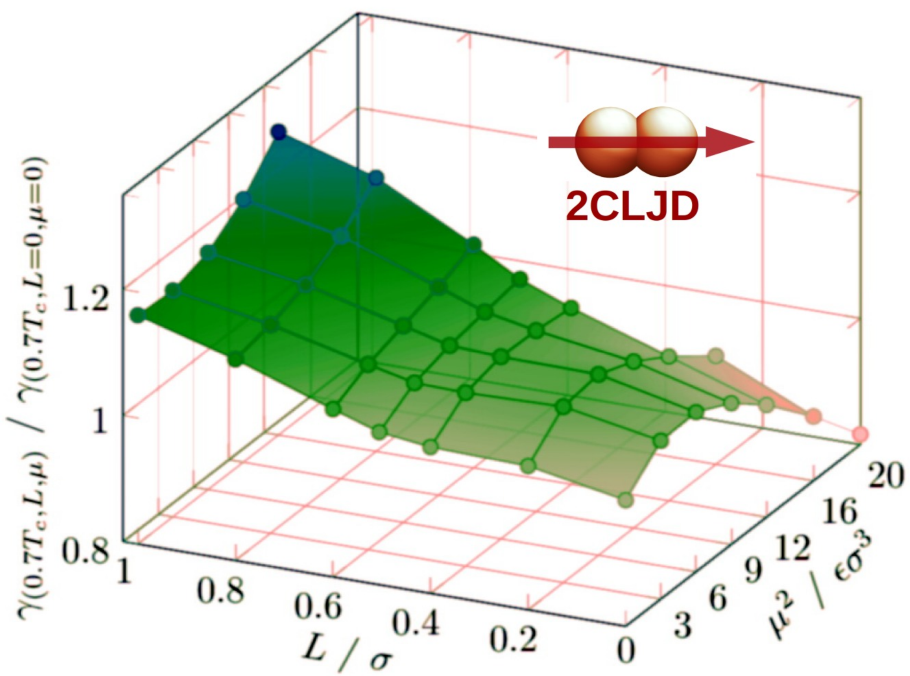
[1]S. Werth, K. Stöbener, P. Klein, K. Küfer, M. Horsch, H. Hasse, *Chem. Eng. Sci.* 121, 110 – 117, **2015**.

Reparameterization by multicriteria optmization (MCO)

2 LJ Centres + Quadrupole



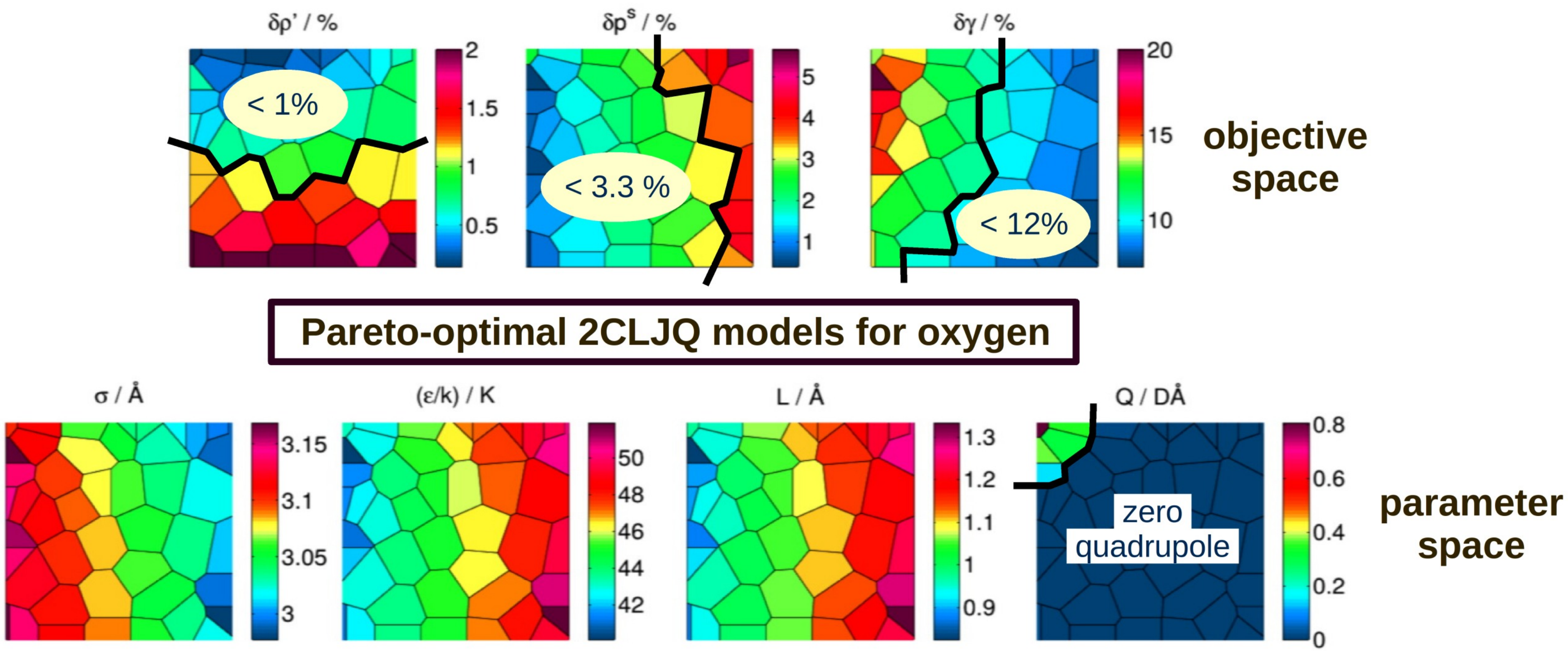
2 LJ Centres + Dipole



- Systematic exploration of the physically relevant part of the model parameter space
- Correlation of the 2LJCQ and 2LJCD surface tension by critical-scaling expressions

Reparameterization by multicriteria optmization (MCO)

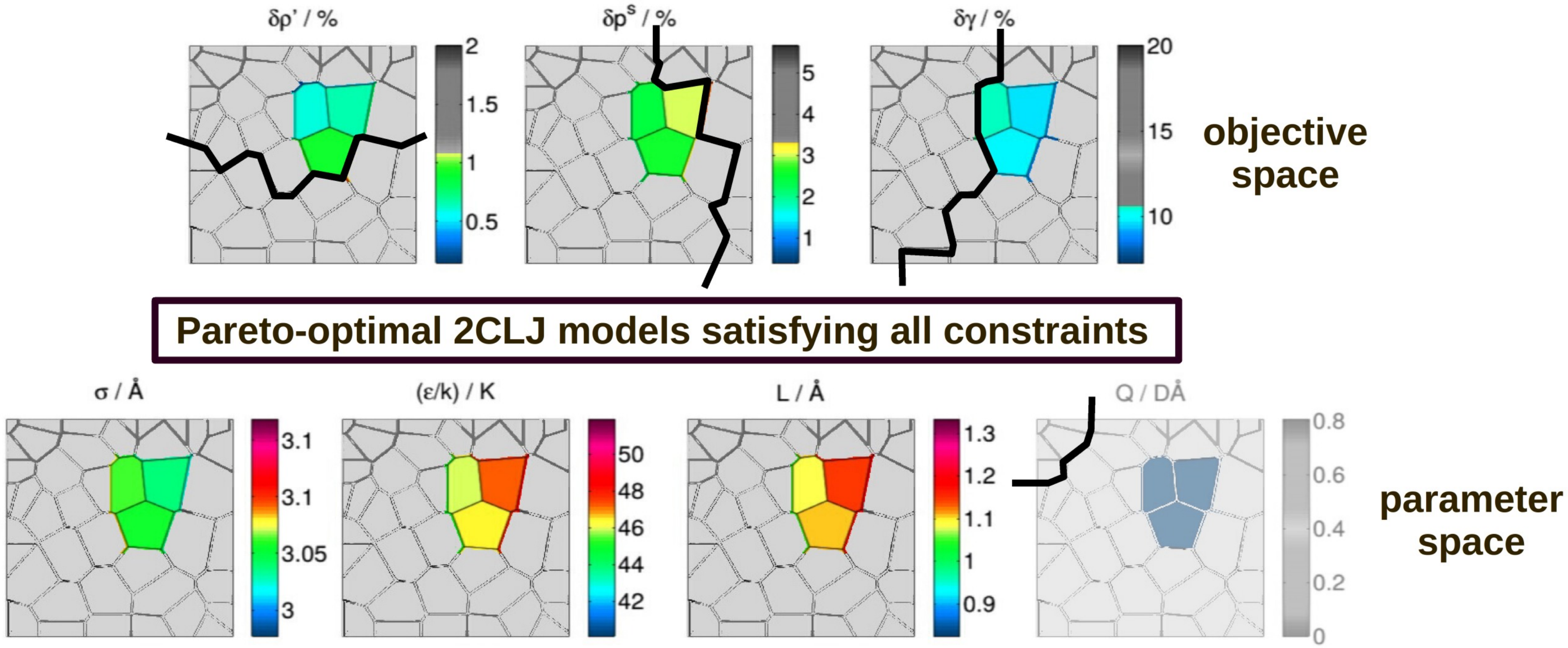
Self-organized patch plots[1] visualizing the Pareto front and the Pareto-optimal models:



[1]K. Stöbener, P. Klein, M. Horsch, K. Küfer, H. Hasse, *Fluid Phase Equilib.* 411, 33 – 42, **2016**.

Reparameterization by multicriteria optmization (MCO)

Self-organized patch plots[1] visualizing the Pareto front and the Pareto-optimal models:



[1]K. Stöbener, P. Klein, M. Horsch, K. Küfer, H. Hasse, *Fluid Phase Equilib.* 411, 33 – 42, **2016**.

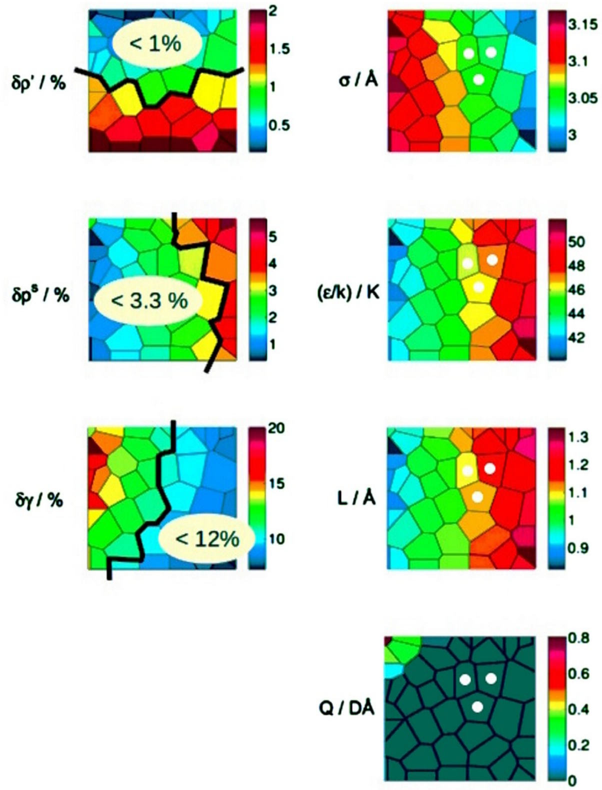
Bespoke model design guided by the end user

“Design your own molecular model in five minutes”

- End-user controlled **multicriteria optimization**, e.g., model selection on patch plots.[1]
- Can be put into practice for **well-characterized model classes** (2CLJQ, Mie-6, …).

known material properties

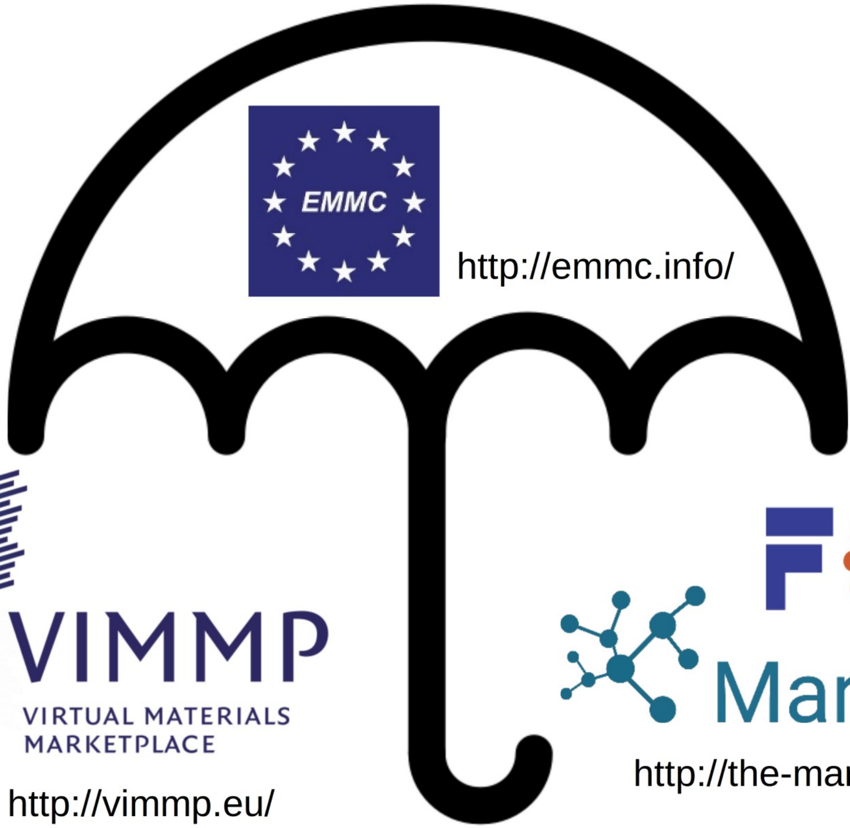
characterized model class properties



bespoke material model by multicriteria optimization

Interoperability in computational molecular engineering

Ambition: European Virtual Marketplace Framework



EMMC

http://emmc.info/

http://the-force-project.eu/

FORCE

MarketPlace

http://the-marketplace-project.eu/

VIMMP

VIRTUAL MATERIALS MARKETPLACE

http://vimmp.eu/

Community-driven materials modelling standardization

Time line of EMMC guided semantic-asset development

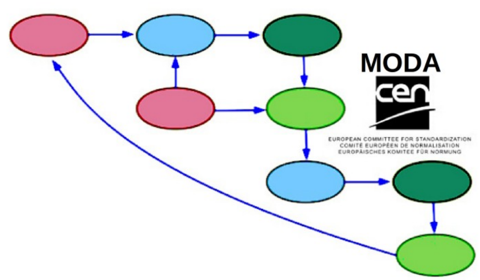
RoMM VI → MODA → CWA 17284 → Ontologies

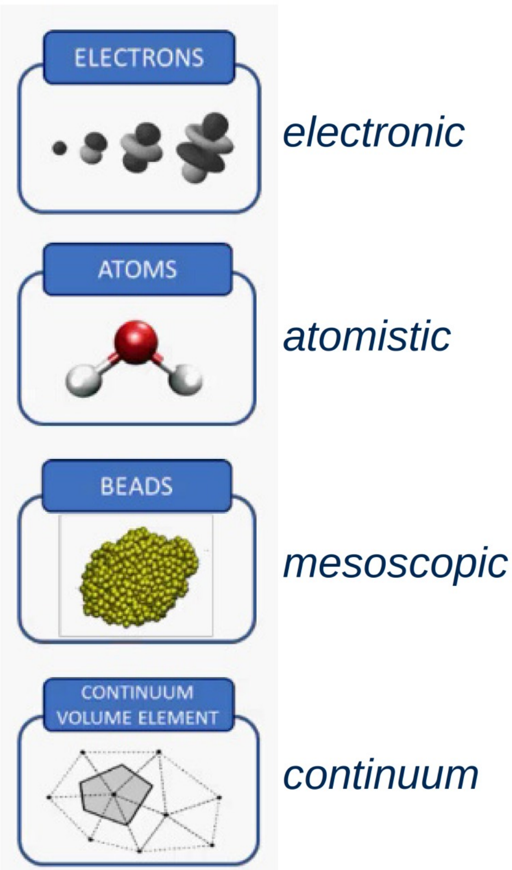
Semi-formalized terminology or vocabulary

MODA workflow graph language

CEN European standard

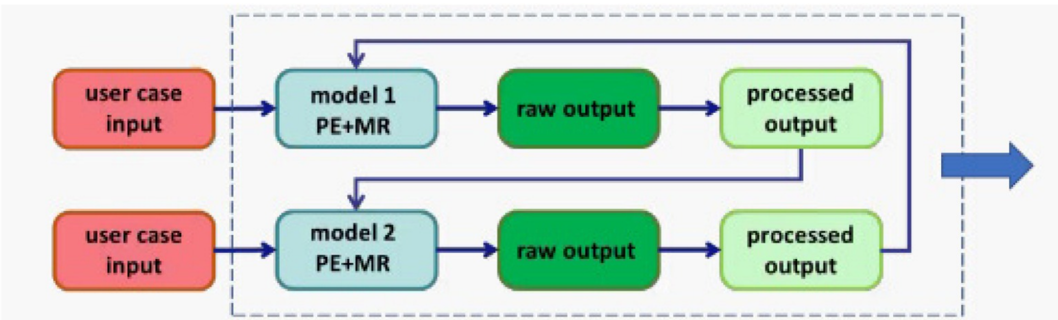
EMMO, EVMPO, marketplace-level, and subdomain-specific ontologies

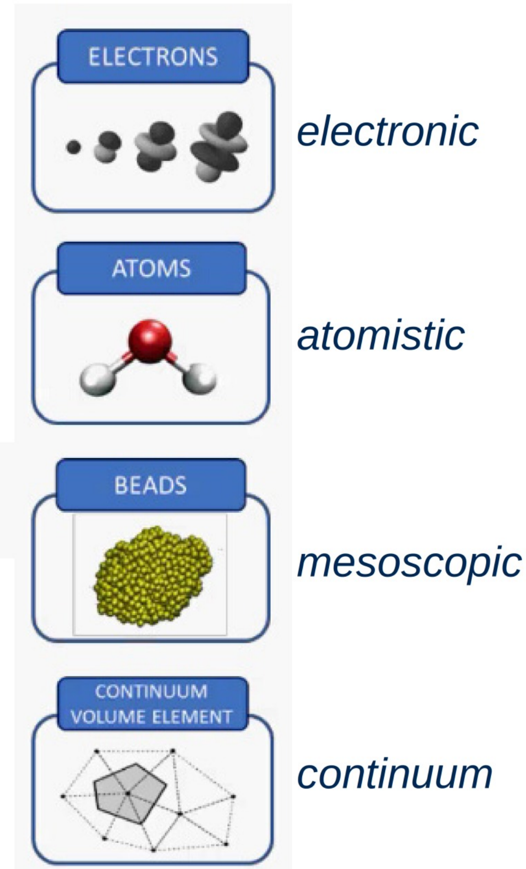


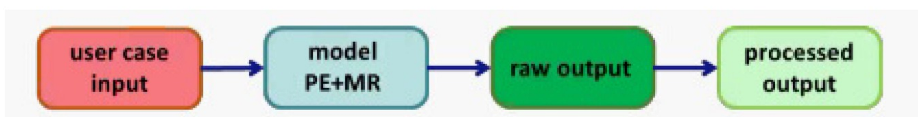


Graph notation for simulation workflow semantics

MODA (“Model data”) workflow graph language (CEN standard by CWA 17284)







MODA

(1) User Case
(2) Model
(3) Solver
(4) Processor

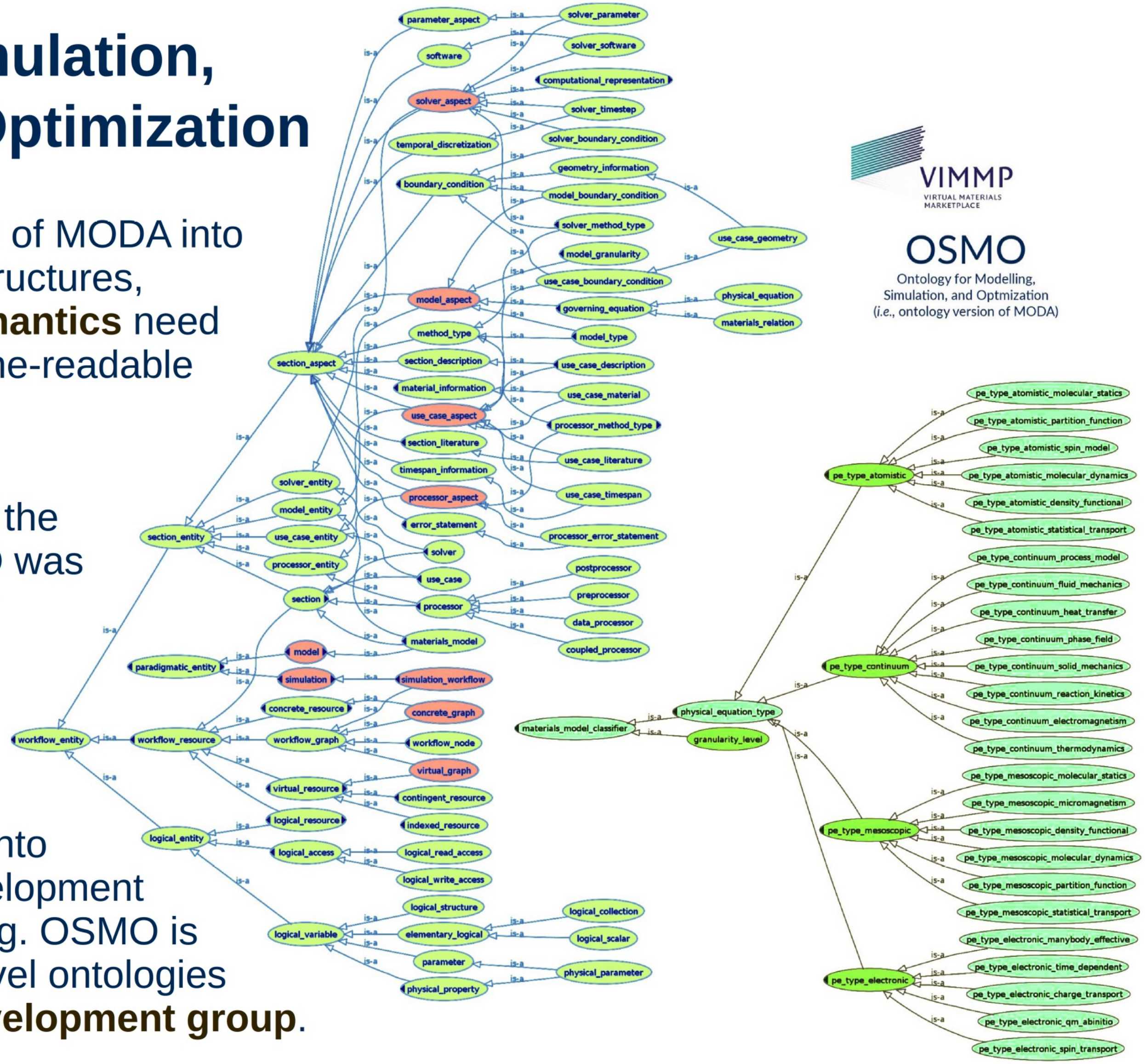
EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

Ontology for Simulation, Modelling, and Optimization

To facilitate the integration of MODA into virtual marketplace infrastructures, **simulation workflow semantics** need to be provided at a machine-readable level of formalization.

Following the approach of the EMMC community, OSMO was developed: **The ontology version of MODA.**

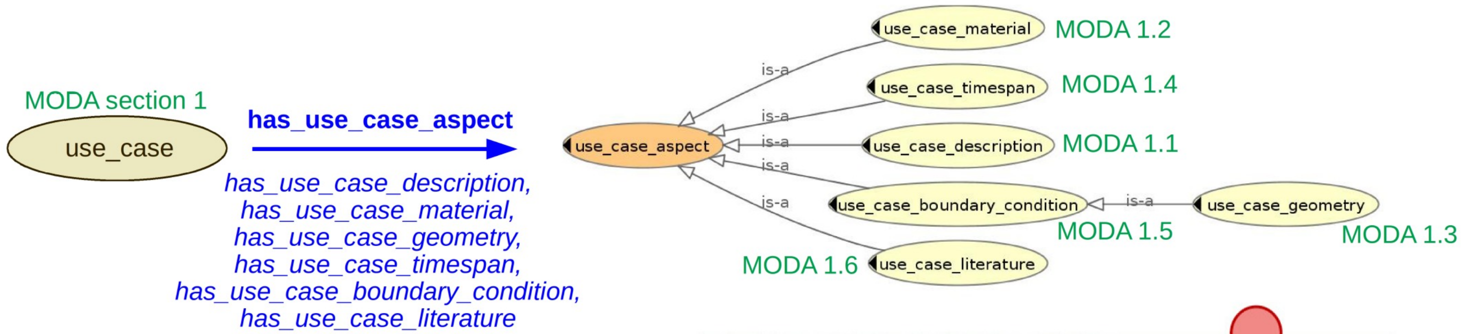
By OSMO, simulation workflow semantics from MODA can be integrated into the ongoing ontology development work in materials modelling. OSMO is one of the marketplace-level ontologies shared in the **EVMPO development group**.



VIMMP VIRTUAL MATERIALS MARKETPLACE

OSMO
Ontology for Modelling, Simulation, and Optmization
(*i.e.*, ontology version of MODA)

Representation of simulation workflow elements



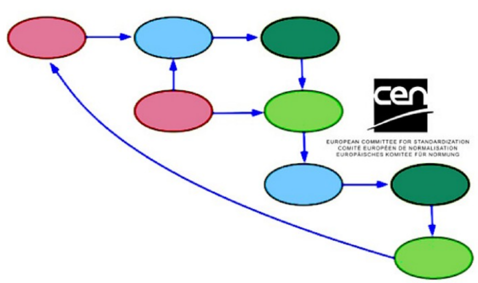
(1) Use Case
(2) Model
(3) Solver
(4) Processor

"sections"

OSMO

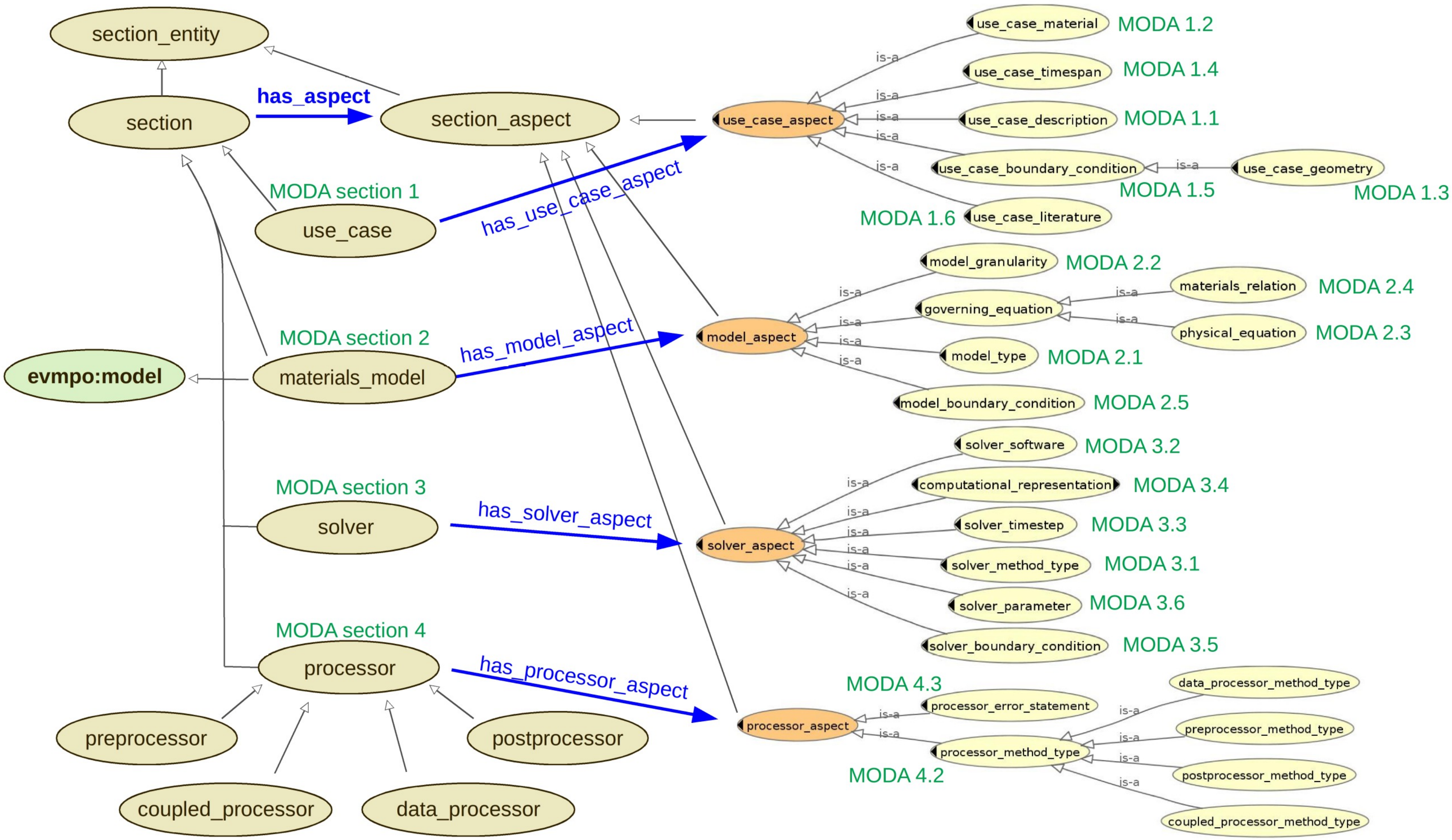
"aspects"

"graphs"

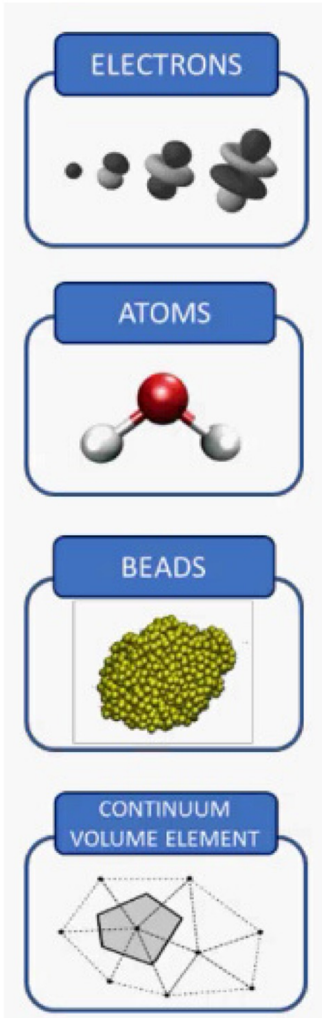


1	Aspect of the User Case/System to be Simulated	
1.1	Aspect of the User Case to be Simulated	*Describe the aspects of the user case textually.* *No modelling information should appear in this box. This case could also be simulated by other models in a benchmarking operation!* *The information in this chapter can be end-user information, measured data, library data etc. It will appear in the pink circle of your workflow picture.* *Simulated input which would have been calculated by another model should not be included (but in chapter 2.4)* *Also the result of pre-processing necessary to translate the user case specifications to values for the physics variables of the entities van be documented here.*
1.2	Material	*Describe the chemical composition. …and the values used for properties and from which database these are taken. If pre-processing was needed please specify the methodology.*
1.3	Geometry	*Size, form, picture of the system (if applicable)* *Note that computational choices like simulation boxes are to be documented in chapter 3.*
1.4	Time Lapse	*Duration of the case to be simulated.* *This is the duration of the situation to be simulated. This is not the same as the computational times to be given in chapter 3.*
1.5	Manufacturing Process or In-Service Conditions	*If relevant, please list the conditions to be simulated (if applicable). These can be boundary, initial and global conditions.* *E.g. heated walls, external pressures and bending forces.* *Please note that these might appear as terms in the PE or as boundary conditions, and this will be documented in the relevant chapters.* *Note: These conditions will be expressed in physics relations in Ch 2.4* *Please specify the values used for parameters and from which database these are taken. If pre-processing was needed please specify the methodology.*
1.6	Publication on this Data	*Publication documenting the simulation with this single model (if available and if not already included in the overall publication).*

Representation of simulation workflow elements



Physical equation taxonomy: Representation in OSMO



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)

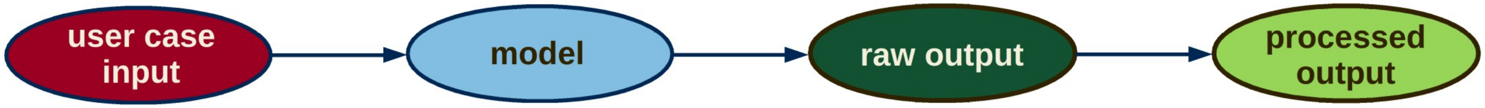
mesoscopic

- 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

Semi-intuitive notation and ambiguities in MODA graphs



"What is the exact meaning of the blue arrows?"

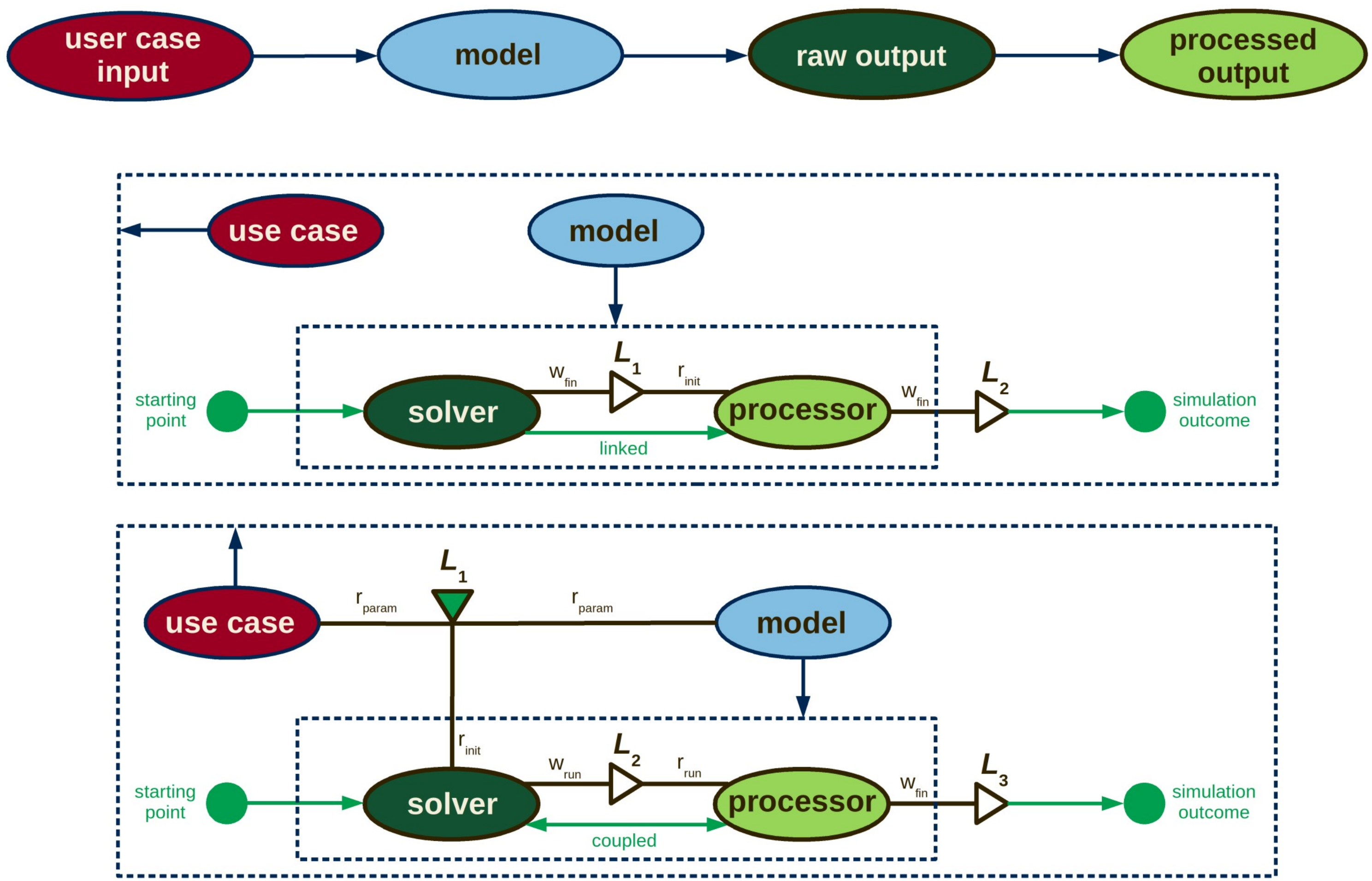
Examples

A **model** can **apply to** a part of the workflow; relation "osmo:applies_to".

The model can be parameterized at workflow execution time; then a MODA arrow may represent **logical data transfer**; n.b., technical data transfer such as file I/O may or may not occur.

Workflows may contain conditional or iterative operations that are active only under certain conditions; in OSMO, such elements are referred to as **virtual graphs**. The relation "osmo:instantiates" relates a concrete graph to a virtual graph.

Graph notation accounting for logical data transfer (LDT)

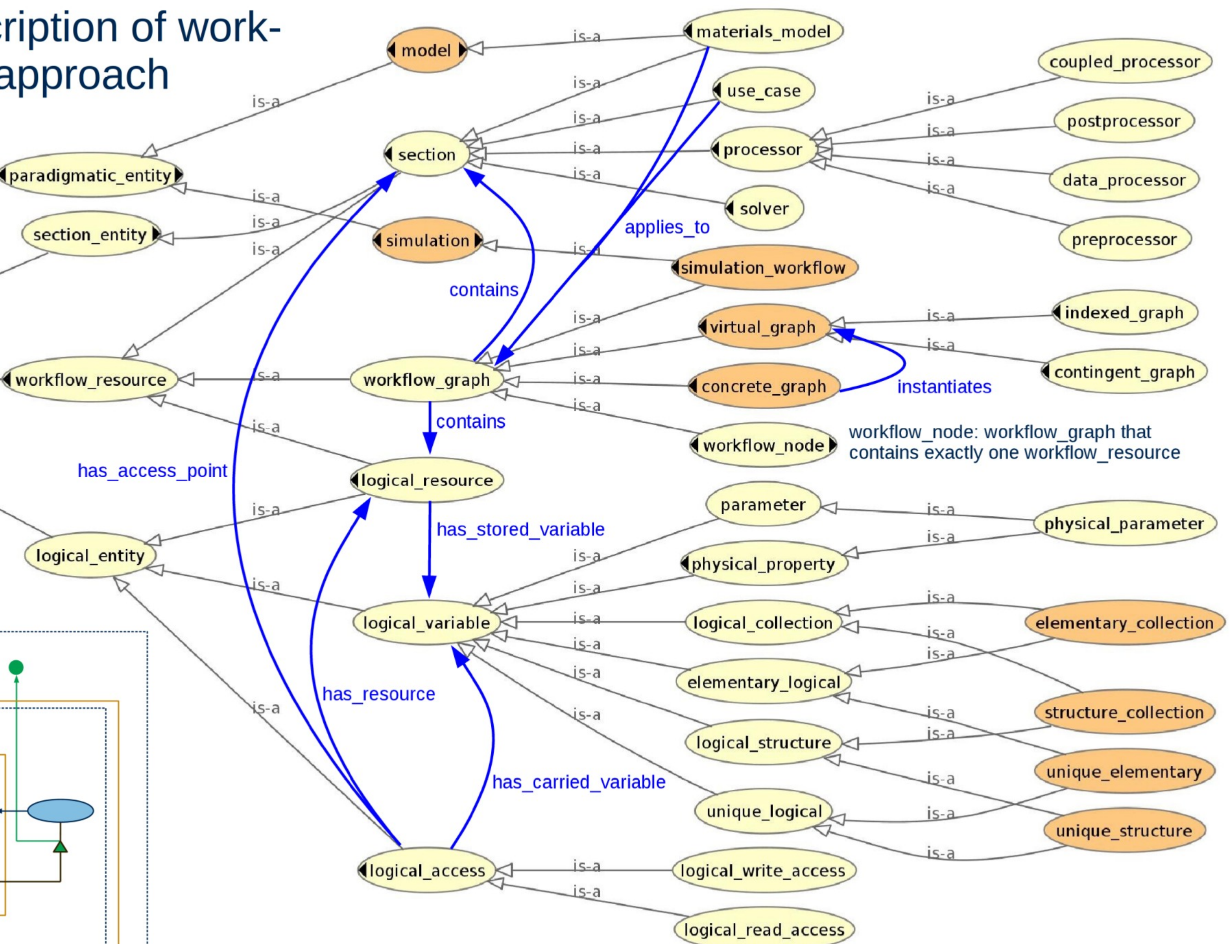


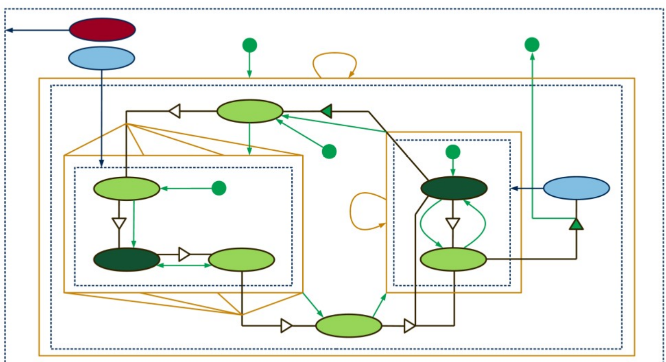
Simulation workflows and logical data transfer in OSMO

Logical (not technical) description of work-flows, following the MODA approach

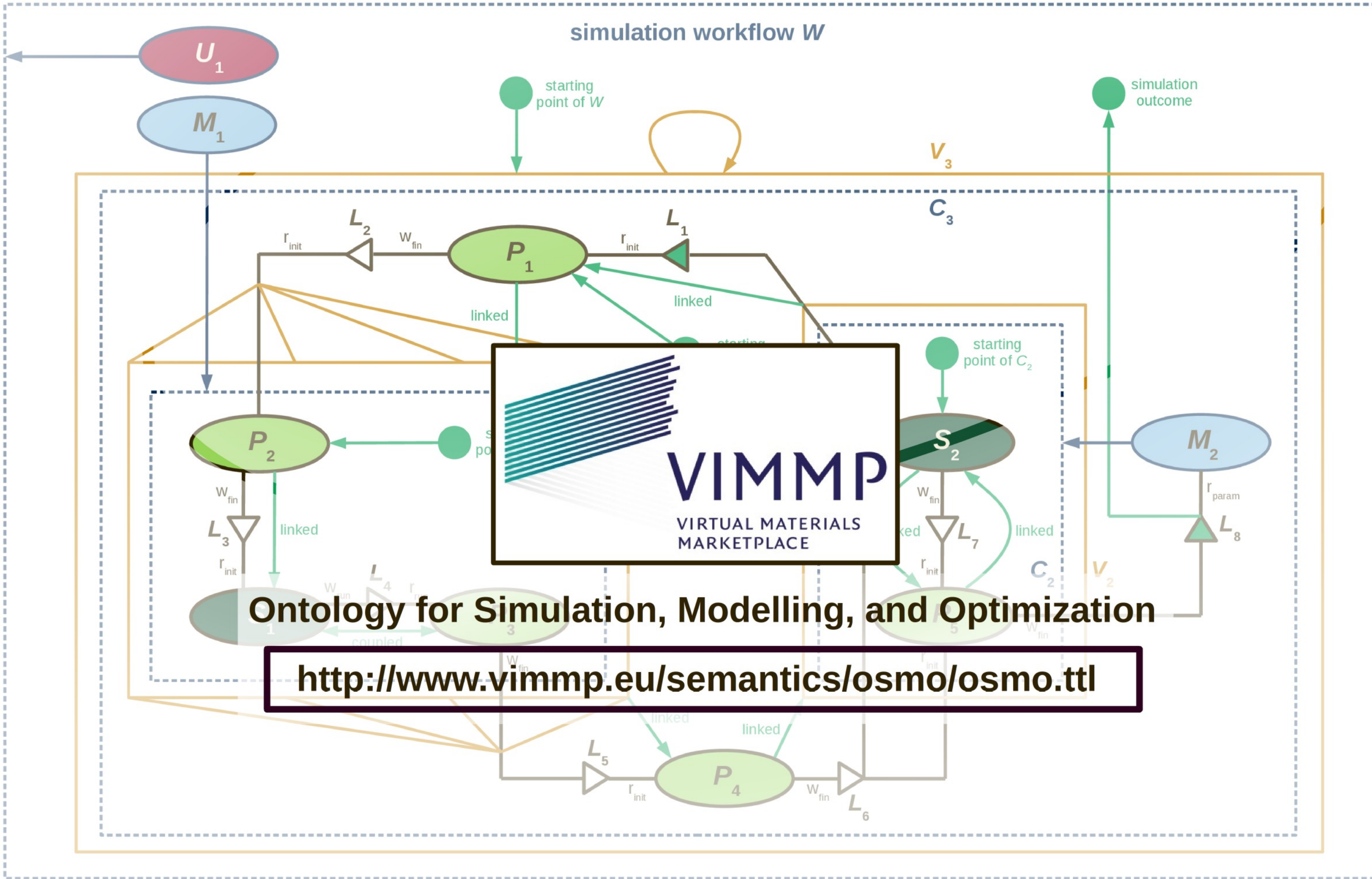
Ontology-level formaliza-tion and less ambiguous representation of the relations between sections

Immediate cor-respondence to section description from MODA





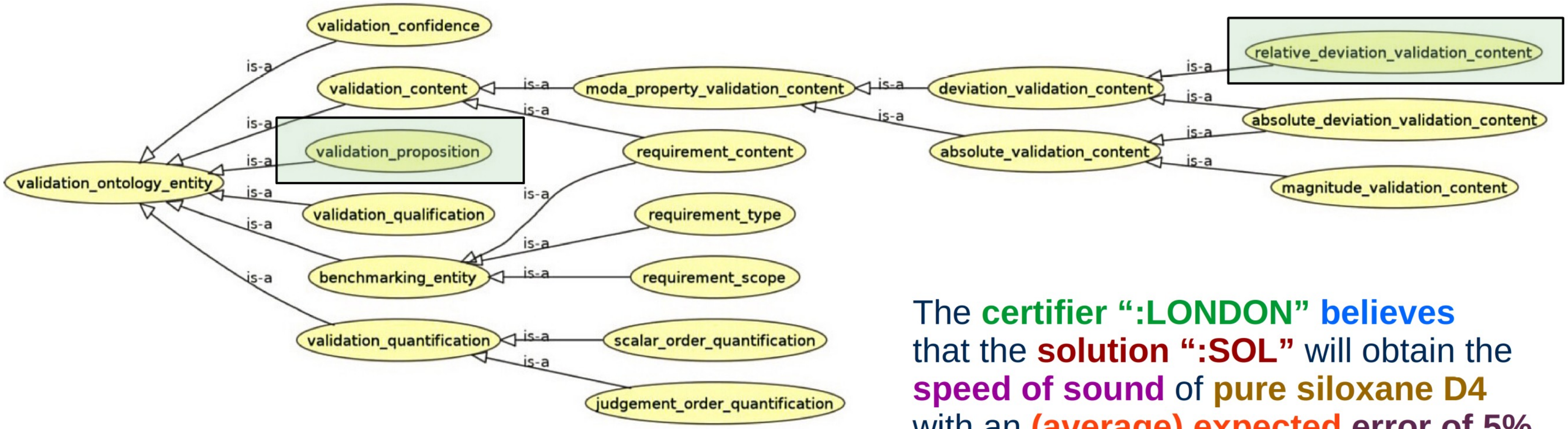
Simulation workflows and logical data transfer in OSMO



Ontology for Simulation, Modelling, and Optimization

http://www.vimmp.eu/semantics/osmo/osmo.ttl

Pragmatic interoperability: Reliabile models and roles



The **certifier ":LONDON" believes** that the **solution ":SOL"** will obtain the **speed of sound** of **pure siloxane D4** with an **(average) expected error of 5%**.

:SOL_ACCURACY a :solution_validation_certificate;
:has_certifier :LONDON;
:refers_to_solution :SOL;
:states **:SOL_ACCURACY_PROP.**

:SOL_ACCURACY_PROP a :validation_proposition;
:has_confidence :STATEMENT_OF_BELIEF;
:has_qualification :EXPECTATION;
:has_content **:SOL_ACCURACY_CONT.**

:SOL_ACCURACY_CONT a :relative_deviation_validation_content;
:asserts_magnitude 0.05;
:refers_to_material :D4_PURE;
:refers_to_property :SPEED_OF_SOUND.

VIMMP
VIRTUAL MATERIALS MARKETPLACE

Competencies and topics in materials modelling

The training ontology will include **topic** and **operator** catalogues.

OTRAS: Ontology for Training Services based on EVMPO, CCSO, IAO, and ICALTZD

mm_topic_basic (codes **1XXX** and **2XXX**):
Basic prerequisites for materials modelling.

mm_topic_computational (codes **3XXX**):
Computational and numerical aspects of materials modelling.

mm_topic_data (codes **4XXX**):
Data science and technology aspects.

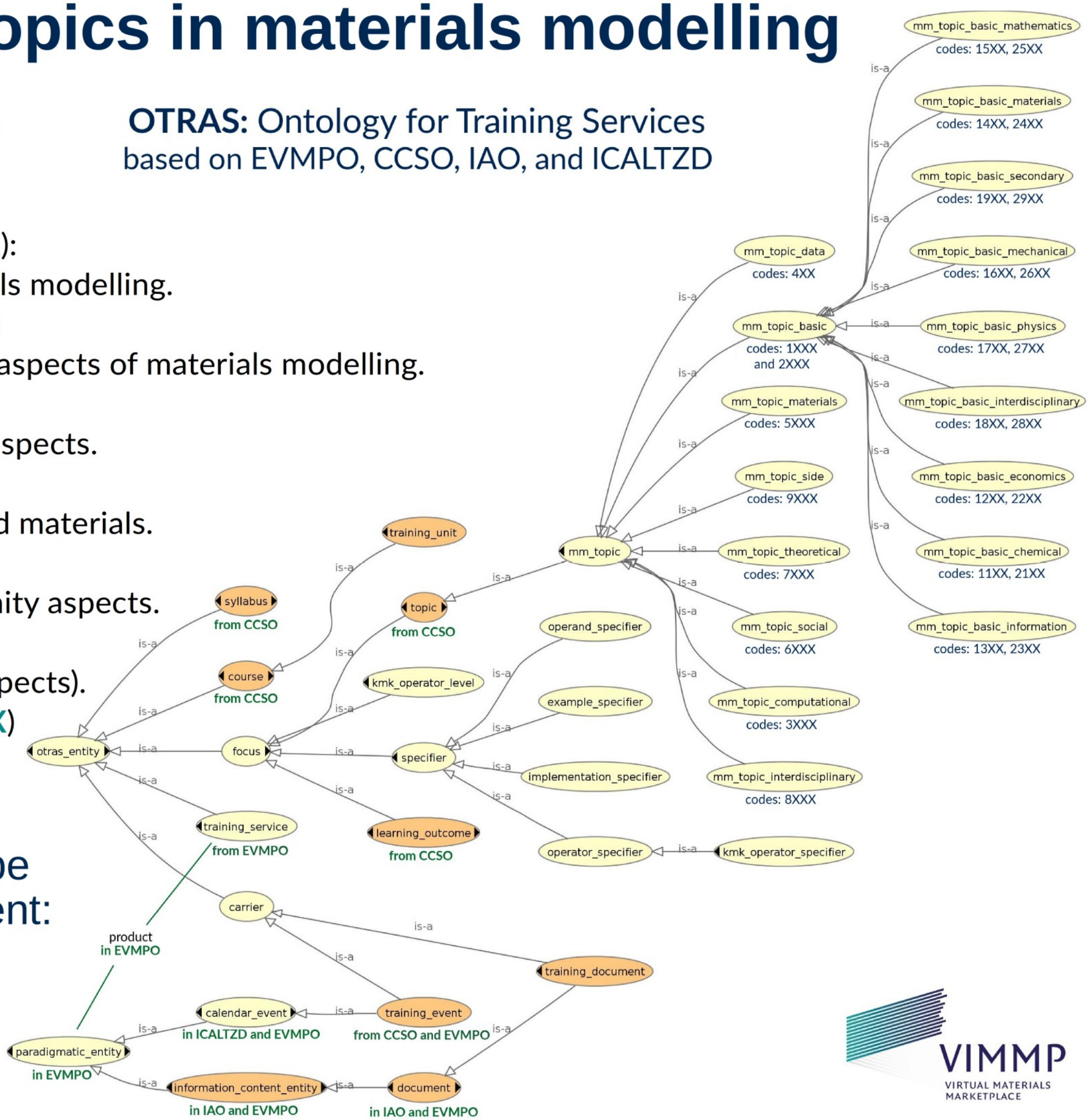
mm_topic_materials (codes **5XXX**):
Topics related to fluid and solid materials.

mm_topic_social (codes **6XXX**):
Social, economic, and community aspects.

mm_topic_theoretical (codes **7XXX**):
Theory (non-computational aspects).

mm_topic_interdisciplinary (codes **8XXX**)

mm_topic_side (codes **9XXX**):
Topics from other disciplines



Pre-existing semantic assets to be considered for further development:

- **ACM Computing Classi-fication System (CCS)**
- **APS Physics Subject Headings (PhysH)**

Operators for learning outcomes (from KMK)

"After successfully completing **X_1**, **participants can X_2** with respect to **X_3** by doing **X_4**; for example, **X_5**." (Note: **X_4** and **X_5** are not required, and **X_1** is not an outcome.)

1XX – Operators for **basic competencies:** "to name/label" (code **120**), "to outline/present" (code **130**), "to list/give" (code **140**), "to write a lab report/data log" (code **150**), "to sketch" (code **160**), "to draw" (code **170**).

2XX – Operators for **intermediate competencies:** "to compare" (code **215**), "to deduce" (code **220**), "to estimate" (code **225**), "to analyse and identify" (code **230**), "to apply" (code **235**), "to calculate" (code **240**), "to describe" (code **245**), "to find" (code **250**), "to explain" (code **255**), "to describe and explain" (code **260**), "to formulate" (code **265**), "to derive" (code **270**), "to sort/group/classify" (code **275**), "to test/verify" (code **280**), "to investigate/examine" (code **285**), "to generalize" (code **290**), "to summarize" (code **295**).

3XX – Operators for **advanced competencies:** "to propose a hypothesis" (code **320**), "to evaluate" (code **330**), "to justify/give reasons" (code **340**), "to comment on/assess" (code **350**), "to prove" (code **360**), "to discuss" (code **370**), "to interpret" (code **380**), "to plan" (code **390**).

Operators for learning outcomes and expert competency

"After successfully completing X_1, **participants can X_2** with respect to X_3 by doing X_4; for example, X_5." (Note: X_4 and X_5 are not required, and X_1 is not an outcome.)

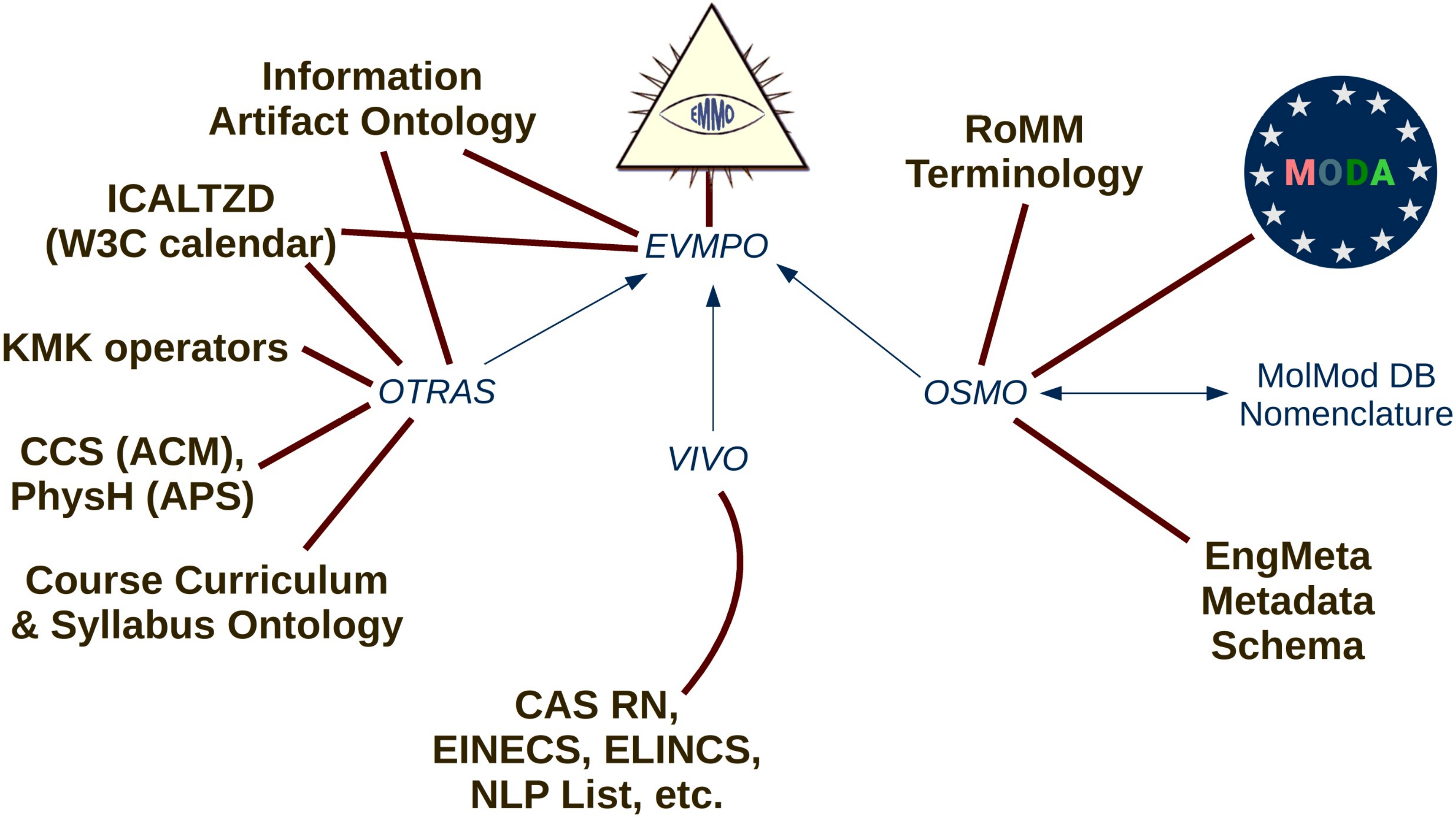
1XX – Operators for **basic competencies:** "to name/label" (code **120**), "to outline/present" (code **130**), "to list/give" (code **140**), "to write a lab report/data log" (code **150**), "to sketch" (code **160**), "to draw" (code **170**).

2XX – Operators for **intermediate competencies:** "to compare" (code **215**), "to deduce" (code **220**), "to estimate" (code **225**), "to analyse and identify" (code **230**), "to apply" (code **235**), "to calculate" (code **240**), …, "to summarize" (code **295**).

3XX – Operators for **advanced competencies:** "to propose a hypothesis" (code **320**), "to evaluate" (code **330**), "to justify/give reasons" (code **340**), "to comment on/assess" (code **350**), "to prove" (code **360**), "to discuss" (code **370**), …, "to plan" (code **390**).

4XX – Operators for **expert competencies:** "to review/evaluate critically" (code **420**), "to advise/manage" (code **425**), "to characterize experimentally" (code **430**), "to document" (code **435**), "to carry out professionally" (code **440**), "to correspond" (code **445**), "to teach" (code **450**), "to plan/project" (code **455**), "to conduct an exam" (code **460**), "to systematize" (code **465**), "to expand/extend/generalize" (code **470**), "to simplify/reduce" (code **475**), "to innovate/develop" (code **480**).

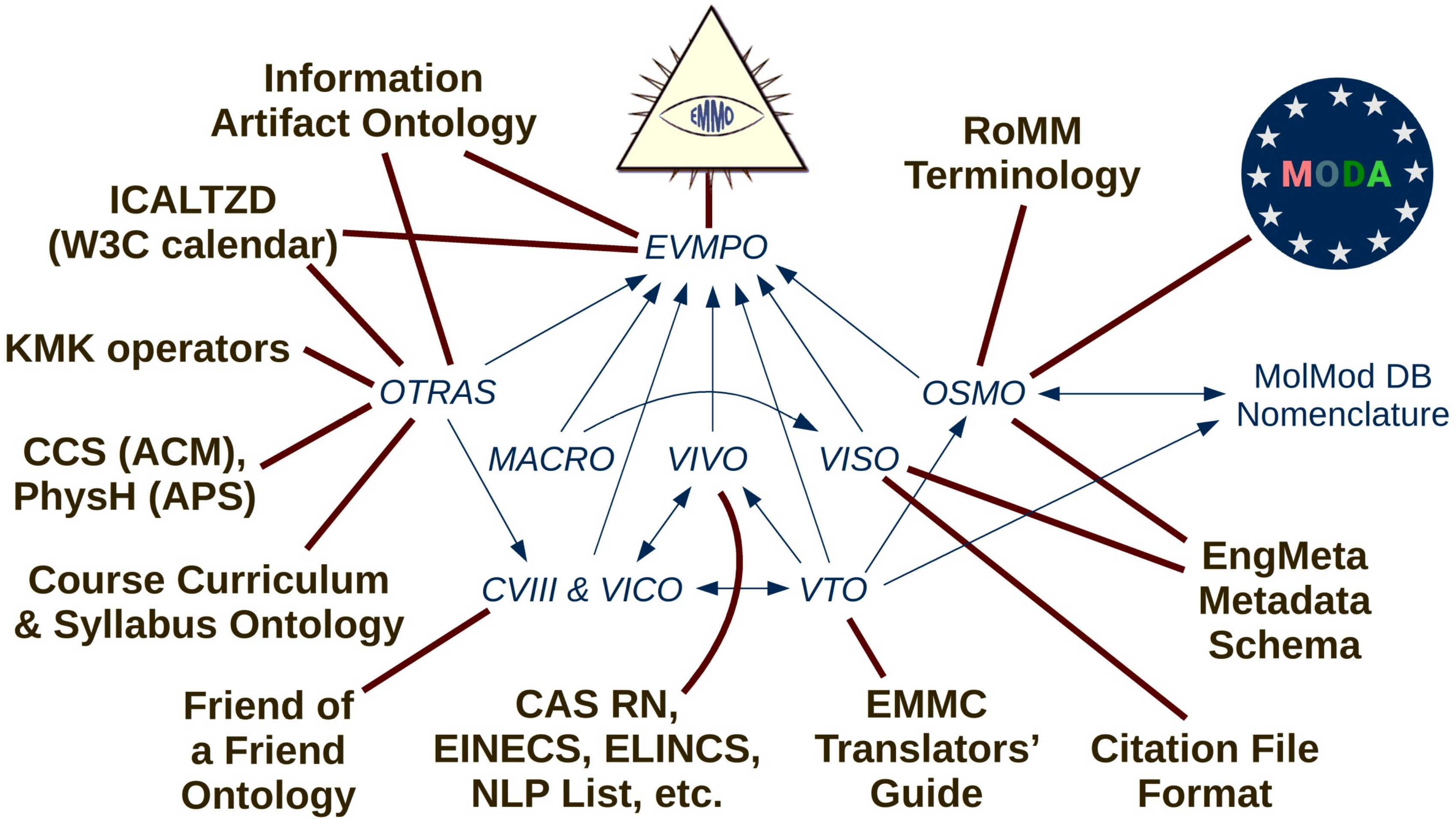
VIMMP ontologies and their relation to existing assets



Blue: Own semantic assets

Black: Related external semantic assets with an overlap or interaction

VIMMP ontologies and their relation to existing assets



Blue: Own semantic assets

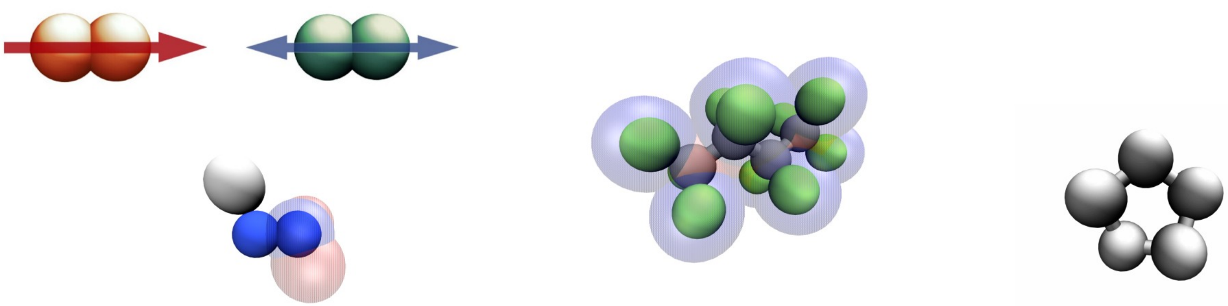
Black: Related external semantic assets with an overlap or interaction

Paradigm shift in materials modelling

Material model design as a <u>specialization</u> practised by expert academics

- A very limited community of people is qualified to design material models;
- for each substance, an expert develops a model, to be looked up from literature;
- it is usually not disclosed how the model was designed and optimized.

Users need a background (e.g., in computational molecular engineering), to assess the model quality, and they **cannot actively contribute** to adjusting the model without repeating the expert work.



Molecular Model Database: MolMod DB

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

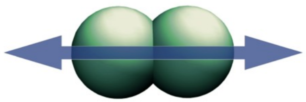
Paradigm shift in materials modelling

Material model design as a <u>specialization</u> practised by expert academics

- A very limited community of people is qualified to design material models;
- for each substance, an expert develops a model, to be looked up from literature;
- it is usually not disclosed how the model was designed and optimized.

Users need a background (e.g., in computational molecular engineering), to assess the model quality, and they **cannot actively contribute** to adjusting the model without repeating the expert work.

Materials modelling as a <u>technology</u> accessible to all industrial engineers



- Models are flexibly adjustable, e.g., by multicriteria optimization, based on a preceding characterization of the model class;
- bespoke models for specific user cases can be obtained with little effort; or automatically, interoperating with other platforms (e.g., process simulation)
- reliable statements on the model uncertainty are available to the user.

Users who may not possess a background in computational molecular engineering are enabled to **actively guide** the material model design, supported by **translators**.

Significant collaboration and interactions acknowledged:

Aachen	–	Youness Bami, Andreas Fiseni, Georg Schmitz
Berlin	–	Jadran Vrabec
Bremen	–	Welchy Leite Cavalcanti, Peter Schiffels
Kaiserslautern	–	Hans Hasse, Peter Klein, Simon Stephan, Katrin Stöbener, Stephan Werth
Stuttgart	–	Christoph Niethammer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 760907.

VIMMP
VIRTUAL MATERIALS MARKETPLACE