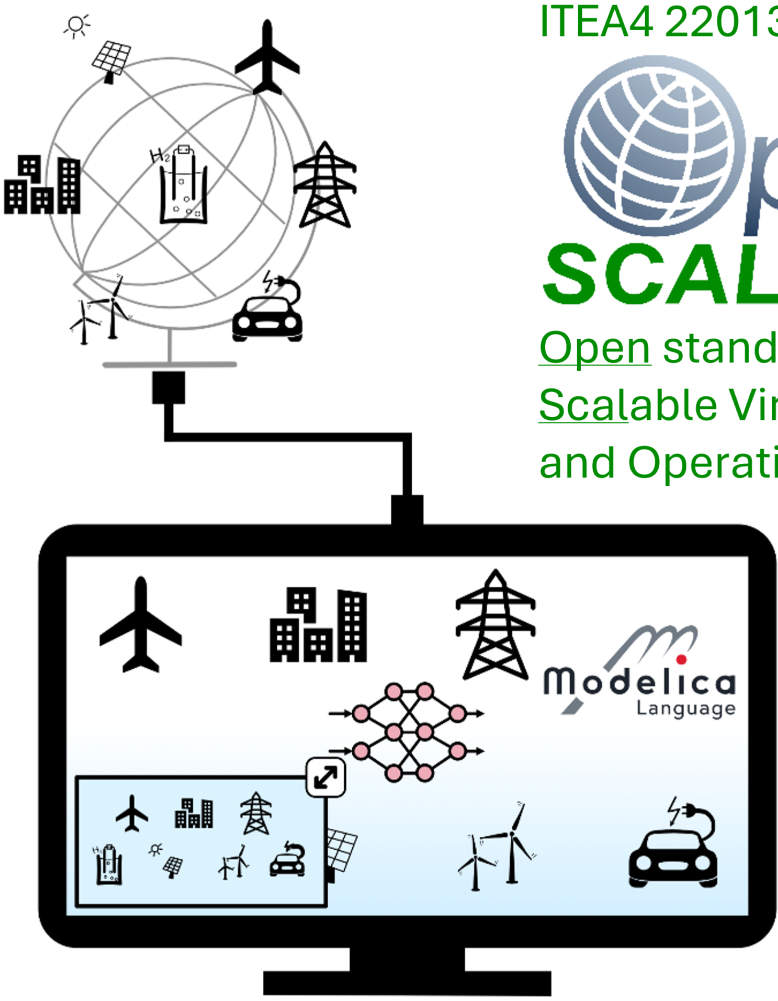

ITEA4 22013

Open standards for
Scalable Virtual Engineering
and Operation

Deliverable D6.3

State-of-the-Art Analysis

Lena Buffoni, Oliver Lenord, Lars Mikelsons, Martin Otter

Version v2, July 28, 2025

Table of Contents

Project Acronyms

Acronym	Description
AI	Artificial Intelligence
COP	Coefficient of Performance
CPS	Cyber Physical System
CSP	Credible Simulation Process
DAE	Differential-algebraic equation
eFMI	Functional Mock-up Interface for embedded systems
FMI	Functional Mock-up Interface (standard for model exchange and co-simulation)
FMI component	A model in FMI format (= FMU)
FMU	Functional Mock-up Unit (= an FMI component)
HVAC	Heating, Ventilation and Air Conditioning
LOTAR	Long Term Archiving and Retrieval
LSS	Large-scale System
M&S	Modelling and Simulation
Modelica	Standard for modelling of cyber-physical systems
MoSSEC	Modelling and simulation information in a collaborative system engineering context
MiL/SiL/HiL	Model/Software/Hardware in the Loop
NeuralODE	Ordinary Differential Equation, where a Neural Network (NN) defines its derivative function
NN	Neural Network
OD	Operational Domain
ODE	Ordinary Differential Equation
ONNX	Open Neural Network Exchange standard
PeN-ODE	Physics enhanced Neural ODE
PINN	Physics-informed Neural Network
SSP	System Structure and Parameterization (standard for connected FMUs)
UQ	Uncertainty Quantification

Acronym	Description
VV&UQ	Verification, Validation, and Uncertainty Quantification

From Full Project Proposal

ID	Type	Description	Due Month	Access
D6.3	Doc	Update of the ITEA "Living Innovation Roadmap" - will be iteratively delivered, e.g. with each PPR	recur.	Public

1. Modelling standards and software for digital twins

OpenSCALING will improve simulation-based processes and scalable digital twins. One essential part are extensions of the following existing modelling standards that are in wide-spread use for the development of digital twins and are utilized to describe and to exchange dynamic multi-domain models, in particular from the mechanical, electrical, thermal, fluid, control, energy, building, automotive, aerospace domains. The underlying mathematical description are differential, algebraic and discrete equations:

- The Modelica language standard defines an open object-oriented language with 2-dim. object diagrams to model complex, dynamic systems on a high level supporting acausal connections of components defined by first principle equations. This standard is developed since 1997, is supported by > 10 tools, and is in widespread industrial use. Modelica tools support export of causal Modelica models as FMI components (see next item). A large class of advanced Modelica libraries has been developed in the ITEA EUROSYSLIB project. Developments towards decarbonized energy systems for buildings, district energy systems and factories are often performed with the large, open source Modelica Buildings library.
- FMI (Functional Mock-up Interface) is the leading, open standard to exchange dynamic models on a low level using a combination of (a) an XML-File to define the interface of a parameterized input/output block, (b) a dynamic link library to define the executable code that is accessed via a C-API, and (c) other resources all packed together in a zip-file. This standard was developed in the ITEA MODELISAR project and was afterwards further improved ([Blo2012], [Jun2021]). It is supported by > 180 tools and plays a key role in many industries for collaborative workflows and comprehensive cross-domain system level analysis, optimization and virtual tests.
- The SSP (System Structure & Parametrization) open standard is used to define complete systems consisting of one or more connected FMI components including their parameterizations. It is supported by nearly 10 tools. Developments performed in the SetLevel project extend SSP with quality assessment information to support a Credible Simulation Process Framework.
- eFMI (Functional Mock-up Interface for embedded systems) is a recent open standard intended as exchange format for workflows and tool chains from physical models to embedded production code. An eFMI component is FMI compliant and can therefore be simulated by FMI tools to perform Software-in-the-Loop testing. Utilizing an eFMI component on an embedded device requires however dedicated tool support for eFMI. This standard was developed in the ITEA EMPHYSIS project ([Len2021]). The first tools with eFMI support are currently coming to the market.

- The ISO 10303-243:2021 MoSSEC (modelling and simulation information in a collaborative system engineering context) standard is an industrial effort to make progress in the representation of the elements “that together comprise a set of "results" for a study including the audit-trail of what is to be done, and what has been done, and evolution”, enabling “the representation of the definitions of models and key values that are part of the modelling” among others to allow the proper reuse of simulation models in a collaborative system engineering environment. Currently, some initial reference implementations can be found in the standard. LOTAR (LOng Term Archiving and Retrieval) is an international consortium with the prime objective to create and deploy the EN/NAS 9300 series of standards for long-term archiving and retrieval of digital data in the aerospace domain. The LOTAR MBSE workgroup suggests the usage of Modelica, FMI and SSP as a basis ([Coi2021]).

Other important tools for simulation-based processes and digital twins:

- MATLAB and Simulink to design, simulate and deploy input/output blocks and especially controllers.
- Simscape to model and simulate multi-domain physical systems.
- JuliaSim (rebranded as Dyad) to model and simulate multi-domain physical systems within the Julia ecosystem. The available model libraries are currently very limited when compared with Modelica or Simulink/Simscape. Advantage is the easy combination with many open-source Julia packages, e.g., for error propagation or machine learning.
- Open source packages from the Julia ecosystem such as ModelingToolkit.jl or Modia.jl provide high-level descriptions of multi-domain models, however lack a graphical user interface. The available model libraries are currently very limited when compared with Modelica or Simulink/Simscape. Advantage is the easy combination with many open-source Julia packages, e.g., for error propagation or machine learning.

References

- [Blo2012]: Blochwitz, Otter, Akesson, Arnold, Clauß, Elmqvist, Friedrich, Junghanns, Mauss, Neumerkel, Olsson, Viel (2012): Functional Mock-up Interface 2.0: The Standard for Tool independent Exchange of Simulation Models. DOI 10.3384/ecp12076173.
- [Jun2021]: Junghanns, Blochwitz, Bertsch, Sommer, Wernersson, Pillekeit, Zacharias, Blesken, Mai, Schuch, Schulze, Gomes, Najahfi (2021): The Functional Mock-up Interface 3.0 - New Features Enabling New Applications. DOI 10.3384/ecp2118117.
- [Len2021]: Lenord, Otter, Bürger, Hussmann, Le Bihan, Niere, Pfeiffer, Reicherdt, Werther (2021): eFMI: An open standard for physical models in embedded software. DOI 10.3384/ecp2118157)

- [Coi2021]: Coïc, Murton, Mendo, Williams, Tummescheit, Woodham (2021): Modelica, FMI and SSP for LOTAR of Analytical mBSE models: First Implementation and Feedback. DOI 10.3384/ecp2118149.

2. Multi Physics Simulation

OpenSCALING addresses the field of system simulation which is characterized by the interaction of sub-models from different physical domains. Current industrial trends result in new challenges for multi-physics simulations:

- Automotive: HVAC-systems of electric vehicles do not provide a comfort function for the passengers only. Keeping the battery system at the right temperature level is essential for a correct function and a long battery lifetime. In contrast to combustion engines which produce enough waste heat, electrical vehicles must generate extra heat which directly reduces the range. Heat pumps and very efficient HVAC-systems are countermeasures. This leads to more complex and larger multi-physics system models, which are challenges for simulation tools, and to higher requirements to the accuracy of models, which increases the modelling effort and requires expert knowledge.
- Buildings and Energy: Also, in these fields a trend to larger simulation models with more subsystems coming from different physical domains can be observed. The energy field has to consider the interactions between different kinds of energy production like fossil, wind and solar energy. State of the art buildings have multivalent sources for electrical energy, heating, and cooling which are coordinated by energy management systems. To assess the efficiency of such systems, simulations over the whole year are necessary to consider the seasonal weather effects. Simulation tools are confronted with large-scale models which have to be computed extremely fast.
- All industries: The necessity to include sub-models from different domains leads to an increased exchange of simulation models via FMI. The concept of Terminals, introduced with FMI 3.0 in 2022, simplifies the error-proof interconnection of FMUs due to bus and physical connectors for the modeller. But the FMI 3.0 Terminals concept is based on causal connectors. The selection, which signal becomes an input or an output still needs to be negotiated between the involved parties. Inappropriate constellations lead to algebraic loops over large parts of the combined models which often leads to numerical problems and/or a reduction of the computational performance. The OpenSCALING innovation regarding acausal FMU-interconnections will significantly improve this situation.

3. Artificial intelligence for modelling and simulation

Following the rapid progress in the field of machine learning in computer vision, classification and further in the last decade, industrialization of these approaches and methodologies already happened or is ongoing. The topic of hybrid modeling, the use of machine learning as part of common simulation models, was initially raised in the 90s ([Psi1992], [Hon1992], [Joh1992]), however, has not been able to gain industrial relevance over the last decades. Progress in the field of AI hardware, as well as the further development and adaptation of algorithms for sensitivity analysis, have allowed this field of research to flourish once again today. Important contemporary works in this field where Neural ODEs in 2018 ([Che2018]) and physics-informed neural networks (PINNs) in 2019 ([Rai2018]).

In classical (engineering) modeling, the right-hand side of an ordinary differential equation (ODE) is generally represented as a symbolic expression derived from physical equations or mechanistic relationships. A Neural ODE is an ordinary differential equation, where a neural network defines the right-hand-side. If the right-hand-side is a combination of neural networks and physical equations no term has become established yet. Within OpenSCALING the term PeN-ODEs (Physics enhanced Neural ODEs) is used since these models can be trained with the same methods as Neural ODEs, however contain explicit physical formulations. In this way, for example, previously constant parameters can be replaced by NNs that can in principle depend on arbitrary other quantities. Another way to utilize NNs is to improve the state derivatives computed using physical equations to match measurement data before passing them to the ODE solver. The gradients with respect to the NN parameters, required for training, can either be computed using AD (Automatic Differentiation) through the solver or other sensitivity analysis methods for ODEs. In contrary, for PINNs the physical equations are used as a regularization term in the loss and the model itself is a NN without any physical equations. This difference can be in turn used to classify different approaches to integrate physical knowledge into hybrid models. Typically, using physical equations inside the model leads to a better extrapolation capability, especially in presence of (time-dependent) inputs. However, if low computational effort is favored often models without physical equations are employed. In recent years many approaches for generating such **surrogates** have been developed. However, method development, testing and validation is in most of the cases done using academic examples. Within project PHyMoS selected approaches, e.g., Proper Orthogonal Decomposition and MeshGraphNets, are investigated in industrial use cases. While great potential has been proven, upscaling to LSS results in unacceptable training times if no application specific measures are taken. After showing the potential of hybrid modelling in academia in recent years, this approach was evaluated in first use cases among different domains from climate modelling, simulation of the human cardiovascular system ([Thu2021]) and modelling of fluid flows to driving simulation. Within the ITEA project UPSIM accuracy boosts up to 40% on validation data were shown for the hybrid modelling of a brake system as PeN-ODEs in form of a Neural FMU and comparable results for PeN-ODEs of a vehicle's vertical dynamics ([Thu2022]). Current restrictions of FMI hinder to develop and train more sophisticated architectures for Neural FMUs that would enable upscaling hybrid modelling with FMUs to more complex systems. Vice versa for the integration of NNs into Modelica models there is no standard allowing a seamless integration in the according system simulation standards. Currently, open standards like NNEF (Neural Network Exchange Format) or ONNX (Open Neural Network Exchange) lack either tool support, sufficient big community, required feature set or suitability for usage in embedded systems. This is in turn also a challenge for the integration of fast surrogates generated with AI methods, e.g., using the aforementioned methods from PHyMoS, into a larger system context.

Moreover, recently, the question of uncertainty quantification (UQ) was also raised for hybrid modelling ([Psa2022]), however, an according methodology or toolset available for industry is missing.

To conclude, it is shown that using AI in modelling has the potential to become the standard approach for complex systems. What is still missing for a quicker and wider adoption is on the one hand better support by tools and standards, e.g., extending FMI to represent PeN-ODEs instead of ODEs and provide an interface for relevant methods like sensitivity analysis that is required in order to efficiently train new hybrid models. On the other hand, adoption of UQ methods for according models respectively training methods that can handle large-scale systems are required to apply the available technology for credible models of complex systems.

References

- [Psi1992]: Psichogios, Dimitris C. and Ungar, Lyle H. (1992): "A hybrid neural network-first principles approach to process modeling." AIChE Journal. DOI 10.1002/aic.690381003.
- [Hon1992]: Hong-Te Su and N. Bhat and P.A. Minderman and T.J. McAvoy. (1992): "Integrating Neural Networks with First Principles Models for Dynamic Modeling." IFAC Proceedings Volumes 25. DOI 10.1016/S1474-6670(17)51013-7.
- [Joh1992]: Johansen, Tor A. and Foss, Bjarne A. (1992): "Representing and Learning Unmodeled Dynamics with Neural Network Memories." 1992 American Control Conference. DOI 10.23919/ACC.1992.4792705.
- [Che2018]: Chen, Ricky TQ, et al. (2018): "Neural ordinary differential equations."

Advances in neural information processing systems. DOI 10.48550/arXiv.1806.07366

- [Rai2018]: Raissi, Maziar, Paris Perdikaris, and George E. Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. DOI 10.1016/j.jcp.2018.10.045.
- [Thu2021]: Thummerer, Tobias and Tintenherr, Johannes and Mikelsons, Lars. (2021) "Hybrid modeling of the human cardiovascular system using NeuralFMUs". Journal of Physics: Conference Series. DOI 10.1088/1742-6596/2090/1/012155.
- [Thu2022]: Thummerer, Tobias and Tintenherr, Johannes and Mikelsons, Lars. (2022) "NeuralFMU: Presenting a Workflow for Integrating Hybrid NeuralODEs into Real-World Applications". Electronics. DOI 10.3390/electronics11193202.

4. Credible digital twins

In the ITEA UPSIM project several elements for credible digital twins are developed, such as the Credibility Development Kit. Partially, the Credible Simulation Process Framework from the SetLevel project is utilized that integrates simulation with SSP models into the development and quality assessment of automated driving functions. In both projects emphasis is on the management process to develop credible digital twins. There is a huge literature on other aspects of credible models, such as calibration, verification, validation, uncertainty analysis, sensitivity analysis, Monte Carlo Simulation, Design of Experiments etc… ([Law2019], [NAS2019], [Rie2021]). All these methods are typically not integrated in a modelling software. For example, the uncertainty information is usually defined in the tools that perform uncertainty analysis, and not in the models where the information naturally belongs to.

References

- [Psa2022]: Psaros, Apostolos F., et al. "Uncertainty quantification in scientific machine learning: Methods, metrics, and comparisons." arXiv preprint arXiv:2201.07766 (2022). DOI 10.48550/arXiv.2201.07766
- [Law2019]: Law (2019): How to build valid and credible simulation models. DOI: WSC40007.2019.9004789.
- [NAS2019]: NASA 2019: NASA Handbook for Models and Simulations.
- [Rie2021]: Riedmaier, Danquah, Schick, Diermeyer (2021): Unified Framework and Survey for Model Verification, Validation and Uncertainty Quantification. DOI 10.1007/s11831-020-09473-7.

5. Uncertainty Quantification

5.1. Uncertainty Quantification (UQ) for Model Credibility and Standards-Centric Workflows

The modern view of model credibility builds on the Verification, Validation, and Uncertainty Quantification (VV&UQ) framework: models must be verified for numerical correctness, validated against data, and presented with quantified uncertainty to support trustworthy decisions ([Roy2011]).

A recent survey across multiple industries makes the case for a unified VV&UQ view that covers different application areas, from finite-element models to cyber-physical systems, while keeping credibility evidence portable. Riedmaier et al. propose a modular framework that explicitly links VV&UQ activities to decision-making, highlighting the need for traceable artifacts and reusable evidence throughout a model's lifecycle ([Riedmaier2021]).

Two key developments now define state-of-the-art implementations:

- **Machine-interoperable traceability & process frameworks:** Extension mechanisms, such as layered approaches in FMI and the SSP Traceability standard, allow credibility information to be packaged with models in a consistent way. In addition, continuous credibility processes introduce checkpoints such as traceability, quality assurance, and review points, so that evidence of credibility can be built up and reviewed over time ([Ahmann2022], [Heinkel2021]). OD-aware traceability of model verification and intended use further supports automated relevance checks and scalable reuse in large simulation environments ([Rosenlund2025]).
- **Credibility-aware modeling and data stewardship.** On the modeling side, Otter et al. demonstrate how Modelica models can be extended with traceability, parameter uncertainty, and calibration metadata using the open Credibility library, bringing credibility information closer to the model itself ([Otter2022]). On the data side, the FAIR principles (Findable, Accessible, Interoperable, Reusable) have become a central reference for managing simulation inputs, datasets, and results, with an emphasis on machine-actionable metadata that supports reuse and automation ([Wilkinson2016]).

One of the important focus areas is predictive capability: measuring whether a model can be trusted across its operational domain (OD), not only at validated points. Two main groups of metrics are used. Coverage-based metrics (for example, a modified nearest-neighbor metric) assess how well validation experiments cover the OD, discouraging results outside the tested range and helping to plan effective experiments (

[Atamturktur2015]). Information-theoretic metrics, based on Shannon entropy and Kullback–Leibler divergence, compare distributions from simulations and experiments ([Shannon1949], [Kullback1951]). Recent aerospace studies combine both approaches to deliver OD-aware assessments of predictive capability based directly on validation experiments ([Hallqvist2023]).

In summary, the state of the art is converging on: (i) standards-centric packaging of UQ and credibility data (via layered standard extensions) to enable cross-tool reuse; (ii) process frameworks that integrate credibility into day-to-day simulation workflows; (iii) objective predictive-capability metrics that explicitly consider coverage and information content across the OD; and (iv) FAIR data stewardship and model-embedded credibility metadata to make the entire workflow reproducible and machine-actionable ([Roy2011], [Riedmaier2021], [Otter2022], [Wilkinson2016]).

References

- [Roy2011]: Roy, Christopher J., and William L. Oberkampf (2011): “A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing.” **Computer Methods in Applied Mechanics and Engineering**, 200, 2131–2144. DOI 10.1016/j.cma.2011.03.016.
- [Riedmaier2021]: Riedmaier, Stefan, Benedikt Danquah, Bernhard Schick, and Frank Diermeyer (2021): “Unified Framework and Survey for Model Verification, Validation and Uncertainty Quantification.” **Archives of Computational Methods in Engineering**, 28(4), 2655–2688. DOI 10.1007/s11831-020-09473-7.
- [Ahmann2022]: Ahmann, Maurizio, Van-Thanh Le, Frank Eichenseer, et al. (2022): “Towards Continuous Simulation Credibility Assessment.” **Proceedings of the Asian Modelica Conference 2022**, 171–182. DOI 10.3384/ecp193171.
- [Heinkel2021]: Heinkel, H. M., and K. Steinkirchner (2021): **Credible Simulation Process – Simulation-based Engineering and Testing of Automated Driving.** Link.
- [Rosenlund2025]: Rosenlund, Erik, Robert Hällqvist, Robert Braun, and Petter Krus (2025): “Objectively Defined Intended Uses, a Prerequisite to Efficient MBSE.” **Proceedings of the American Modelica Conference 2024**, 29–42. DOI 10.3384/ecp20729.
- [Otter2022]: Otter, Martin, Matthias Reiner, Jakub Tobolář, Leo Gall, and Matthias Schäfer (2022): “Towards Modelica Models with Credibility Information.” **Electronics**, 11(17):2728. DOI 10.3390/electronics11172728.
- [Wilkinson2016]: Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. (2016): “The FAIR Guiding Principles for scientific data management and stewardship.” **Scientific Data**, 3:160018. DOI 10.1038/sdata.2016.18.

- [Atamturktur2015]: Atamturktur, Sez, Matthew C. Egeberg, François M. Hemez, and Garrison N. Stevens (2015): “Defining coverage of an operational domain using a modified nearest-neighbor metric.” **Mechanical Systems and Signal Processing**, 50–51, 349–361. DOI 10.1016/j.ymssp.2014.05.040.
- [Shannon1949]: Shannon, C. E. (1949): “A Mathematical Theory of Communications.” **The Bell System Technical Journal**, 27(3), 379–423.
- [Kullback1951]: Kullback, S., and R. A. Leibler (1951): “On Information and Sufficiency.” **Annals of Mathematical Statistics**, 22(1), 79–86. DOI 10.1214/aoms/1177729694.
- [Hallqvist2023]: Hällqvist, Robert, Erik Rosenlund, Per Norström, et al. (2023): “Toward Objective Assessment of Simulation Predictive Capability.” **Journal of Aerospace Information Systems**, 20(3), 152–167. DOI 10.2514/1.I011153.