Innovationszentrum für Prozessdatentechnik e.V.

Technical report no. 2022–C
original version

Topology-based construction of business-integrated material modelling workflows

---

Date: 19th July 2022

Authors: Preisig, H. A.; Klein, P.; Konchakova, N.; Hagelien, T. F.; Friis, J.; Horsch, M. T.

Dissemination:

- In *Proc. ESCAPE-32*, Elsevier (ISBN 978-0-32395879-0), 2022.

Funding information:

- EC, H2020, MarketPlace, GA no. 760173
- EC, H2020, VIMMP, GA no. 760907
- EC, H2020, VIPCOAT, GA no. 952903
- RCN/Forskningsrådet, Bio4Fuels, project no. 257622

Accessibility:

- doi:10.5281/zenodo.6861461
- https://zenodo.org/communities/inprodat/

PROCEEDINGS OF THE 32nd European Symposium on Computer Aided Process Engineering
(ESCAPE 32), June 12-15th, 2022, Toulouse, France.
L. Montastruc, S. Negny (Editors)

Topology-Based Construction of Business-Integrated Material Modelling Workflows

Heinz A. Preisig[a], Peter Klein[b], Natalia Konchakova[c], Thomas F. Hagelien[d], Jesper Friis[e] and Martin T. Horsch[f, g, h]

[a]*Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway*
[b]*Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, Germany*
[c]*Institute of Surface Science, Helmholtz-Zentrum Hereon, Geesthacht, Germany*
[d]*Climate and Environment Department, SINTEF Ocean, Trondheim, Norway*
[e]*Materials and Nanotechnology Department, SINTEF Industry, Trondheim, Norway*
[f]*STFC Daresbury Laboratory, UK Research and Innovation, Daresbury, UK*
[g]*School of Psychology and Computer Science, University of Central Lancashire, Preston, UK*
[h]*Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway*
heinz.preisig@chemeng.ntnu.no

Abstract

Designing a new product requires information from the business and physical domains, which implies integrating business decision tools with process and material simulation processes to form an overall workflow. The integration involves coupling the business workflow management systems with analysis tools, optimization and decision support systems, which require process simulations and an integrated data transfer service. The process simulation, in turn, will in general model multiple layers of time scales and thus is also in need of data transfer between different solvers. Here we discuss the main components in the light of a coating-design project.

Keywords: Molecular and multiscale modelling, applied ontology, process data technology.

1. Project Background

The Horizon 2020 project VIPCOAT constructs an innovation platform for new active protective coatings based on materials modelling and optimization. The project's application is coatings for the aerospace industry, and its objective is to introduce novel approaches for corrosion protection by active inhibiting pigments as nano-additives. VIPCOAT is to support end-users in developing new and effective corrosion barriers for metal surfaces, deploying environmentally friendly technologies, and, in parallel, provide a decision support system for business integration. VIPCOAT implements a multi-layer digital structure to enable open innovation for production processes and the design of value-added product chains. VIPCOAT relies on the concept of a Pareto chain along a value-adding B2B2B (business-to-business-to-business) sequence enabling collaborative, transparent decision processes. The approach builds on multicriteria optimization (MCO) and implements a set of decision tools that allow exploring individual Pareto fronts interactively.

Usually, coatings contain several components. Primarily the coating matrix is the material that makes up the coating itself. In addition, several additives are mixed in, protecting from UV, giving colour, and inhibiting corrosion in coating defects. When producing a paint, one has several objectives: purpose-related performance, the toxicity of the involved materials, and costs of production and application. Typically, these objectives are not possible to satisfy simultaneously. Also, with having different components, usually, several suppliers are part of the production process. Therefore, it is natural that each member of the suppliers and customers has another set of objectives. And each member is affected by all of those being close to them in the process. Consequently, if one has to define *optimality*, one is confronted with multiple, incompatible business-related objectives or KPIs (Key Performance Indicators). Hence, overall, one deals with two levels of multi-objective optimisation problems: the lower one is the product, the upper the multi-player business layer. The project will create a digital platform for coating formulation, development and optimization, which could serve as a computational coating marketplace.

The VIPCOAT platform will comprise a collection of physics and data-based materials models, a data space, and an associated ontology-driven service that enables a smooth and simple data transfer between Business Decision Support Systems (BDSS), and the MoDeNa software orchestrator. For this purpose, the team develops new semantic network-based technologies and approaches to ensure an automatic data exchange for multi-scale simulations and multi-level material and business software components. These semantic technology developments are aligned with the Elementary Multiperspective Material Ontology (EMMO), cf. Francisco Morgado et al. (2020), in line with a series of efforts coordinated through activities of the European Materials Modelling Council (EMMC ASBL) and the Innovation Centre for Process Data Technology (Inprodat e.V.).

2. Business-to-Business process level

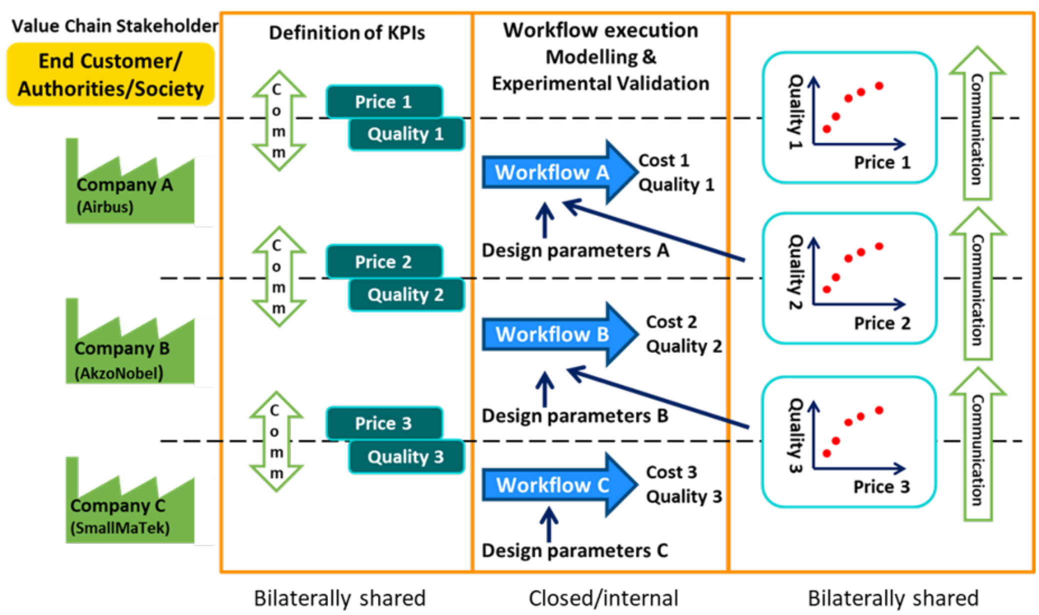

Figure 1: Business-to-Business-to-Business approach implemented in VIPCOAT. Bilateral communication starts the processes, followed by business internal developments and ending in an open access to Pareto-optimal product variants to be used in decision making upstream the production chain.

The VIPCOAT approach supports collaborative decision making along with production and/or value-added chains, using a conceptually new idea: the Pareto chain along Business-to-Business-to-Business (B2B2B) value-added chains as pictorially presented in Figure 1. The value-added chain goes from the bottom to the top. On the bottom, in VIPCOAT SmallMaTek (SMT) producing inhibiting pigments, Company C has some production capabilities associated with costs. As high quality and low cost are almost always contradicting goals, the capabilities to deliver products upstream the value chain are best described using a Pareto front of the best possible compromises. For a company B downstream, in VIPCOAT represented by AkzoNobel, the best possible compromises form a sub-space of the Companies B opportunities and thus enter their design space for fabrication for company A, in VIPCOAT represented by Airbus. Therefore, we have a B2B2B environment in which collaborative decision making on top of transparent decision processes will make a lot of sense. In VIPCOAT, this is directly supported by the concept of a Pareto chain of interactive decision tools, designed as interactive explorers of the individual Pareto fronts. One Business executes a development and/or production process, supported by materials modelling to be described next, internally. Thus the internal cost structure is kept confidential and the transparent part of the B2B2B relation is the Pareto front of prices and associated product variants.

3. Business internal process level

One of the project's main objectives is to reduce the number of experiments, or in other words, replace experiments with predictions based on a physical model of the experimental setup. The first step is thus to design and discuss a model topology using the minimal graphical language we defined (Preisig (2021, 2014)). The topology captures all the fundamental structural assumptions. It represents the physical process as a network of primitive capacities (base entities, control volumes), a minimal set of extensive quantities and a set of mechanisms on how entities interact. Since it is a hierarchical representation, it also shows the assumed mereology of the model. The topology reflects all the main assumptions: what parts of the system are considered, which ones exhibit capacity effects, if they show significant gradients in the intensive quantities, if they are constant or dynamic or event-dynamic or not, how they link up to other capacities and exchange extensive quantity, what exchange is considered relevant, the exchange mechanism, and the resource environment. It should be noted, that the topology does not include geometrical information, like the shape of a capacity, also called a control volume.

Our Process Modelling software (ProMo) defines an application-focused ontology that defines the fundamental entities and their mathematical representation. The mereological information is encoded into the model's hierarchical tree. The internal nodes in the tree, we term composite entities. Replacing parts of the topology, usually composite entities, by surrogates generates the skeleton of the workflow. In the case of the coating process, as shown in Fig. 2 it is the two gray boxes that may be replaced by a surrogate solving the input relations in a separate task, a PDE solver like OpenFoam, while the leaching process is solved involving molecular modelling on which we do not expand in this exhibition. When running a workflow, the different tasks do exchange data, in our case this would be reading salinity, but also transferring information about the simulation of the coating and the simulation of the material and its inhibition. The same applies with the interaction of the coating and the leaching process simulations. Since ProMo maps information of the variables and the relations via OWL into a triple store, and the model provides the vectors of information exchange between the blocks, the interoperability problem can be solved using the approach described below. Finally, the task factory generates the simulation code for a solver environment like MatLab or an orchestrator that executes the developed workflow.

The green area in Figure 2 is reflecting only part of the story. In ProMo the access to a data taken from an external database is a controlled process. Figure 3 shows how ProMo maps the variables and equations into an EMMO-extended ontology. Once the model is established, the intertask communications is established over the SOFT/DLite interoperability framework. ProMo maps the variables and equations into an EMMO-extended ontology. Once the model is established, the intertask communications is established over the SOFT-DLite framework.

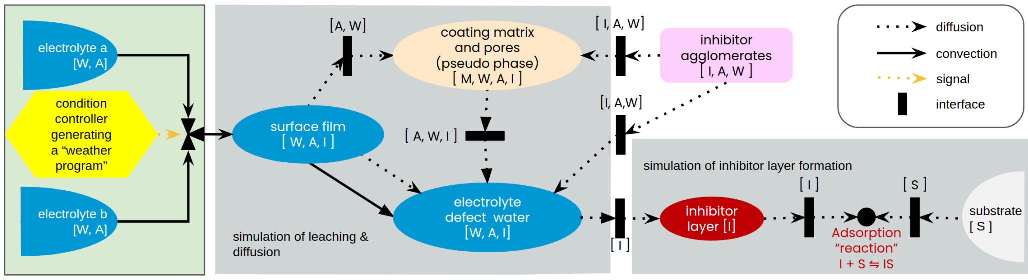

Figure 2: Simplified topology of a coating with inhibitor-loaded nanoparticles. Replacing the two grey boxes by surrogates providing an input/output solution in a separate task, one has the structure of the workflow.

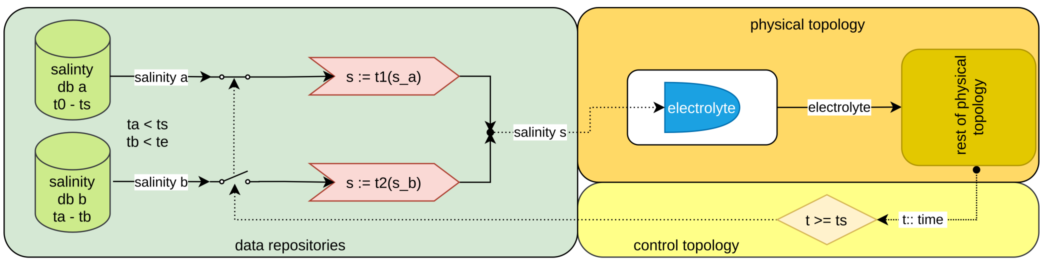

Figure 3: The ProMo view on how the salinity is read from two data sources. The database a covers the first part of the time period, while the database b is used to fill in the remainder.

4. Interoperability

SOFT/DLite (Hagelien et al., 2017) (Mir et al., 2020) represents information exchange between a data source and sink as a pipeline that process output data from the producer into a form that is suitable input to the consumer. Metadata schemas and ontologies capture the description of the data's intent and their meaning (semantics), which can be used in conjunction with specialised data readers for data extraction (Schembera, 2021) (Fig. 4). The next step is to augment high-level semantic artefacts with domain-specific knowledge. EMMO aligned mid-level and domain ontologies can be employed to document concepts such as processes and properties/quantities semantically (Horsch, 2021). Data and metadata need to be represented as knowledge graphs, employing semantic technology standards such as RDF, RDFS, and OWL.

EMMO-aligned mid-level and domain ontologies can also support platform interoperability between VIPCOAT and other infrastructures, particularly those developed within *H2020 NMBP* projects and, in the future, projects supported from the *Horizon Europe CL4 resilience and data* lines of funding. Moreover, the EMMC ASBL focus areas and task groups help coordinate a series of ongoing development efforts in this direction. This technology operates at a comparably high level and requires at least the developers of platforms and tools to be familiar both with ontology-based research data management and with the philosophy underlying the EMMO. In contrast, the second line of development of semantic artefacts, equally endorsed by EMMC ASBL, operates at a lower level and is more accessible to domain experts: Based on the terminology of the Review of Materials Modelling (RoMM), the MODA (Model Data) describe simulation workflows, which permits textual descriptions, and is supported by the OSMO's ontology version of MODA (Horsch et al., 2020, 2021). An alignment between semantic artefacts from these two lines of work is not always straightforward because they rely on knowledge graphs that are differently structured. Previous work by Klein et al. (2021) discusses how graph transformation crosswalks can be used to transpose MODA/OSMO representations of simulation workflows, for which annotation and data ingest are comparably easy, into EMMO-aligned mid-level ontologies that are best suitable for

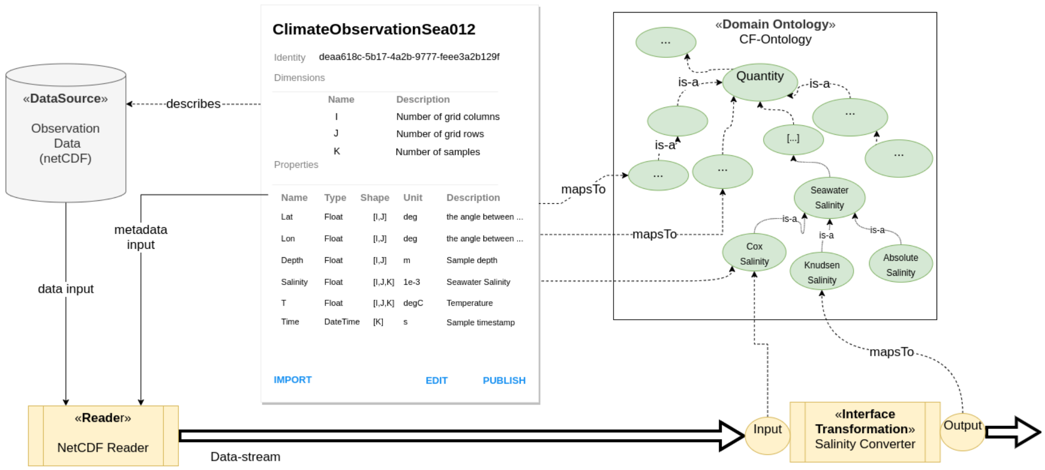

Figure 4: Schematic overview. The key components of the semantic interoperability platform are business data, a business data reader or writer, a metadata representation, domain ontologies, and interface transformations (equations) that connect ontological concepts. The interoperability platform will analyse the use case and give the user options to select a semantic pipeline or manually manage the pipeline pathways.

platform interoperability in line with EMMC ASBL recommendations.

In the present framework, the mapping property *mapsTo* ensures that the semantics of metadata properties are aligned with domain ontology concepts, cf. Fig. 5. Similarly, a data model will represent the schema of application-specific data input on the consumer side. A data ingest system or file format generator using the metadata produces information that the data consumer can interpret. Mapping the metadata schema to the domain ontology concepts captures the semantics of the information. The SOFT/DLite generates the pipeline for the data transfer between the systems, recognising the semantic interpretation and syntactic representations.

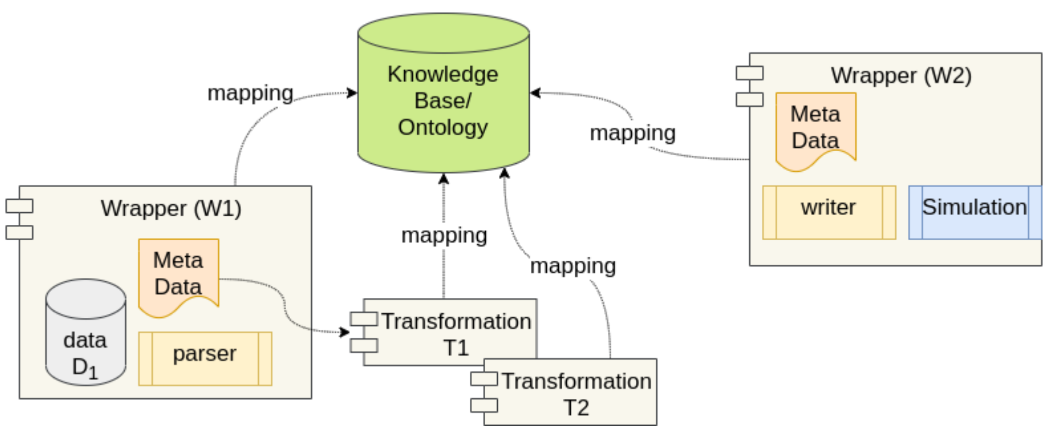

Figure 5: Data resources, data sinks, transformations and simulation packages connect to the knowledge base via mappings between the metadata representation of the business data and domain ontology concepts. Depending on the data model at the receiving side, a pipeline is constructed to transform between schemas for annotating the data; e.g., a pipeline can be constructed from data resource D_1 in wrapper W_1 through the interface transformations (T_1 and T_2) to input for the simulation tool in W_2: $\mathbf{W_1} \rightarrow T_1 \rightarrow T_2 \rightarrow \mathbf{W_2}$.

In the trivial case, the pipeline will only transfer data only. In a more complex scenario, semantic differences between data sources and data sinks require transformations to be taken from an ontology. The ProMo software (see Section 3) is already aligning the equations and variables to EMMO domain ontologies and can therefore be plugged directly into the SOFT/DLite pipelines. Graph-search algorithms find all possible paths that define the pipeline. The end-user is responsible for selecting the most appropriate pathway if multiple exist, closing the connection through mappings and defining new transformations. In addition, the interoperability platform also manages the representation of the information in terms of differences in units (unit systems) and data shape/dimensionality. Unit conversions transformations can usually be inferred automatically, but strategies for managing differences in dimensionality (extrapolation, interpolation, averaging, machine learning-based methods, etc.) needs to be manually determined.

5. Conclusion

We demonstrate a new approach to a complex business-integrated product design on the example of active-protective coatings for the aeroplane industry. We introduce a new approach to interrelate three daisy-chained industries, two producers and one end-consumer. The new business-to-business-to-business environment implements a complex two-level workflow that links the business workflow with the physical simulation workflow, involving all three business partners in forming the final product performance. On the business level, we employ business-process-modelling (BPM) software Camunda and the NTNU's Process Modelling suite (ProMo) to model and generate simulation workflows executed in the MoDeNa platform. The team approaches the data integration of the various software using ontology-based technologies, generating the application interfaces automatically for the different components, providing access to a shared data space and Petri-net based technology for synchronization.

References

J. Francisco Morgado, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, J. Friis, A. de Baas, 2020. Mechanical testing ontology for digital-twins: A roadmap based on EMMO. In: R. García Castro, J. Davies, G. Antoniou, C. Fortuna (Eds.), Proceedings of SeDiT 2020. CEUR-WS, Aachen, p. 3.

T. F. Hagelien, A. Chesnokov, S. T. Johansen, E. A. Meese, B. T. Løvfall, 2017. Soft: a framework for semantic interoperability of scientific software. In: Progress in Applied CFD–CFD2017 Selected papers from 12th International Conference on Computational Fluid Dynamics in the Oil & Gas, Metallurgical and Process Industries. SINTEF akademisk forlag.

M. T. Horsch, 2021. Mereosemiotics: Parts and signs. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOUST 2021). CEUR-WS, Aachen, p. 3.

M. T. Horsch, C. Niethammer, G. Boccardo, P. Carbone, S. Chiacchiera, M. Chiricotto, J. D. Elliott, V. Lobaskin, P. Neumann, P. Schiffels, M. A. Seaton, I. T. Todorov, J. Vrabec, W. L. Cavalcanti, 2020. Semantic interoperability and characterization of data provenance in computational molecular engineering. Journal of Chemical & Engineering Data 65 (3), 1313–1329.

M. T. Horsch, D. Toti, S. Chiacchiera, M. A. Seaton, G. Goldbeck, I. T. Todorov, 2021. OSMO: Ontology for simulation, modelling, and optimization. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOIS 2021 Ontology Showcase). CEUR-WS, Aachen, p. 47.

P. Klein, H. A. Preisig, M. T. Horsch, N. Konchakova, 2021. Application of an ontology based process model construction tool for active protective coatings: Corrosion inhibitor release. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOMI 2021). CEUR-WS, Aachen, p. 26.

Z. M. Mir, J. Friis, T. F. Hagelien, I.-H. Svenum, I. G. Ringdalen, N. Konchakova, M. L. Zheludkevich, D. Höche, jan 2020. Interoperability architecture for bridging computational tools: application to steel corrosion in concrete. Modelling and Simulation in Materials Science and Engineering 28 (2), 025003.
URL `https://doi.org/10.1088/1361-651x/ab6209`

H. A. Preisig, 2014. Visual modelling. Computer Aided Chemical Engineering 34 (ISBN 978-0-4444-63433-7), 730–734, (FOCAPD 8 - Foundation of Computer-Aided Process Design).

H. A. Preisig, 2021. Ontology-based process modelling – with examples of physical topologies. Processes 9 (4), 592.

B. Schembera, 2021. Like a rainbow in the dark: Metadata annotation for HPC applications in the age of dark data. Journal of Supercomputing 77, 8946–8966.

Acknowledgments: Bio4Fuels RCN: 257622, MarketPlace H2020-NMBP-25-2017:760173, VIMMP H2020-NMBP-25-2017:760907, VIPCOAT H2020-NMBP-TO-IND-2020:952903