INPRODAT Innovationszentrum für Prozessdatentechnik e.V.

Technical report no. 2021–G
original version

Application of an ontology based process model construction tool for active protective coatings: Corrosion inhibitor release

---

Date: 9th October 2021

Authors: Klein, P.; Preisig, H. A.; Horsch, M. T.; Konchakova, N.

Dissemination:

- Proceedings of JOWO 2021 (FOMI 2021), paper no. 26

Funding information:

- DFG, project NFDI4Cat, project no. 441926934
- Horizon 2020, project MarketPlace, grant agreement n° 760173
- Horizon 2020, project VIPCOAT, grant agreement n° 952903

Accessibility:

- http://ceur-ws.org/Vol-2969/
- doi: 10.5281/zenodo.5559061
- https://zenodo.org/communities/inprodat/

Application of an Ontology Based Process Model Construction Tool for Active Protective Coatings: Corrosion Inhibitor Release

Peter Klein[1], Heinz A. Preisig[2], Martin Thomas Horsch[3,4] and Natalia Konchakova[5]

[1]*Fraunhofer Institute for Industrial Mathematics, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany*
[2]*Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway*
[3]*High Performance Computing Center Stuttgart, Nobelstr. 19, 70569 Stuttgart, Germany*
[4]*STFC Daresbury Laboratory, UK Research and Innovation, Keckwick Ln, Daresbury WA4 4AD, United Kingdom*
[5]*Institute of Surface Science, Helmholtz-Zentrum Hereon, Max-Planck-Straße 1, 21502 Geesthacht, Germany*

Abstract

Ontology-based integrated materials modelling for an active protective coating system design is presented and applied to a practical example. For this purpose, an ontological methodology implemented using the ProMo (Process Modelling) suite is developed to be used with an open simulation platform (OSP), *i.e.*, a workflow management and orchestration framework that can be integrated into digital infrastructures. The target infrastructures, which are under development in various Horizon 2020 projects, include modelling marketplaces, open innovation platforms, and open translation environments among others. Semantic interoperability for the communication between the involved digital infrastructures, including the simulation hubs, relies on the Review of Materials Modelling (RoMM), MODA (Modelling Data), and the Ontology for Simulation, Modelling, and Optimization (OSMO) in combination with the Physicalistic Interpretation of Modelling and Simulation Interoperability Infrastructure (PIMS-II) mid-level ontology, which is aligned with the Elementary Multiperspective Material Ontology (EMMO) as a top-level ontology. The challenge of addressing semantic heterogeneity is addressed by working toward crosswalks between domain-specific and mid-level ontologies for industrially relevant problems, where knowledge graph transformation is evaluated as a candidate solution for a future implementation strategy. The involved semantic artefacts are platform-agnostic, and their EMMO compliance allows for a specification of executable modelling and simulation workflows on multiple EMMO-compliant OSPs. We demonstrate the presented approach on industrial relevant example for development of active corrosion protection of metallic surfaces.

Keywords

active protective coating, applied ontology, Elementary Multiperspective Material Ontology, graph transformation, molecular modelling and simulation, ontology alignment, process data technology

1. Introduction

Europe has been dedicating substantial efforts in its Framework Programs FP7 and Horizon 2020 to the development of Open Simulation Platforms (OSPs), which consist of model orchestration

FOMI 2021: 11th International Workshop on Formal Ontologies meet Industry, held at JOWO 2021: Episode VII The Bolzano Summer of Knowledge, September 11–18, 2021, Bolzano, Italy
peter.klein@itwm.fraunhofer.de (P. Klein); heinz.preisig@chemeng.ntnu.no (H. A. Preisig); martin.horsch@stfc.ac.uk (M. T. Horsch); natalia.konchakova@hereon.de (N. Konchakova)

tools for the construction of materials modelling workflows and materials modelling and simulation tool repositories. Their harmonization is connected to ongoing work on the Elementary Multiperspective Material Ontology[1] (EMMO) [1, 2] which, as a top-level ontology, supports semantic interoperability, and the MODA (Modeling Data) standard [3] for documenting such modelling workflows. These OSPs implement simulation backbones in various approaches, who follow the idea to make materials modelling widely available, in particular to industrial users, in Horizon 2020 efforts by adding additional services on top of their OSP cores. Among them are business decision support systems, materials modelling marketplaces, open translation environments, open innovation platforms and open innovation environments. All these efforts, at various levels of maturity and rigour, offer aspects of *translation* [4, 5], which in its core deals with the construction of materials modelling workflows fitting to industrial challenges.

Beside translating industrial problems to modelling workflows, translators typically need to implement EMMO-interoperable semantic technology related to the required industrial and simulation workflows. In this paper, we argue that humans providing these translation services are unnecessarily expected to be ontology experts. In our opinion, this is an overburdening of the Translator role that needs to be (semi-)automatized and in this paper, we present such an approach. In previous work, the authors constructed a set of ontology-based tools with the idea to provide a usable abstraction on top of EMMO and MODA supporting interoperability between OSP cores, *cf.* Preisig [6]; in these and similar scenarios, a human in the loop, referred to as the Scientific Data Officer [7], can greatly improve the outcome from the ontological tools developed. Here, these tools are applied in a somewhat different arena: the construction of modelling workflows for industrially relevant applications and the potential to implement automation of modelling workflow construction based on the developed ontological tools.

It turns out that this challenge is feasible, albeit a human inspection and modification of the outcome is still advantageous for constructing optimal modelling workflows.

As a practical example demonstrating the development of ontology based modelling by EMMO extension and MODA application for industrial relevant multi-physics simulations, the release of a corrosion inhibitor in an active protective coating is considered in the current paper.

2. Use Case and Simulation Workflow

2.1. Inhibitor Leaching in Active Protective Coating Processes

It is well known that aggressive environmental exposure and corrosion-induced damage often limit engineering structures' service life. Functional coatings are the most effective and efficient way to protect bridges, off-shore equipment, cars, trains, buildings, ships, aircraft, and daily consumables. The main role of the coating in degradation and damage protection is to provide a dense barrier against corrosive species. However, defects appear in the protective coatings during exploitation of the coated structures opening direct access for aggressive agents to the metallic surface. Including anti-corrosive inhibitor agglomerates into coating polymer matrix provides "smart" corrosion protection when the coating (barrier) is damaged (cracks) by its ability to release a corrosion inhibitor (leaching) which accumulates in the cracks [8]. The

[1]Previously known as the European Materials and Modelling Ontology, with the same abbreviation.

inhibitor, released from the nano-containers, forms an active layer on the substrate surface and thus prevents the corrosion development on the metallic surface [9].

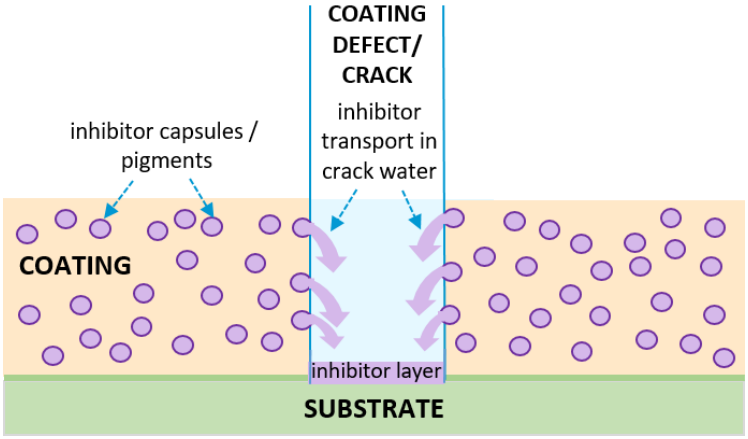

Figure 1: Pigments dissolution: Interpretation of inhibitor leaching processes including model simplification.

Corrosion inhibitor release (leaching), as sketched in Fig. 1, is the main active protective mechanism of corrosion inhibiting primers used in the field of protective coatings [10]. These primers are loaded with sparingly soluble salts to provide a reservoir of corrosion inhibiting ions. In case of a defect, the coating will be exposed to the environment and absorb moisture, which triggers the dissolution of the corrosion inhibitor, the leaching into the damage region (crack), and their agglomeration on top of the metal substrate where inhibitor prevents the surface from damage. The molecular structure of the corrosion inhibitor, the inhibitor leaching rate, and the critical concentration in the damage regeion (crack) are the key factors influencing the immediate and long-term corrosion prevention by this active protective coating process. The development of new coating formulations is a complex process, usually performed by enormous time and resources consuming experimental efforts. An effective way to accelerate novel coatings development is by replacing experimental efforty by materials modelling methods.

The approach summarized above is one of the core scientific ideas of the ongoing H2020 EU project VIPCOAT [11] which is funded under call DT-NMBP-11-2020, “Open Innovation Platform for Materials Modelling.” The main scientific objectives of the project are to establish an Open Innovation Platform for the development of inhibiting active protective coatings and corresponding accelerator tests for assessing their in-service durability, as well as to promote the development of a green technology for active protective coatings based on materials modelling and optimization.

One of the most effective classes of inhibitor agglomerates are Layered Double Hydroxides (LDH) nanocontainers [12]. If LDH nanopigments are present in the corrosive media, a significant reduction of the corrosion rate is observed. LDHs are anion-exchange systems consisting of stacks of positively charged, mixed-metal hydroxide layers, intercalated by layers of anionic species and solvent molecules. These environmentally friendly nano-structures already demonstrated their ability to control the release of active inhibition species under certain

environmental conditions. Their functioning is twofold: not only to release the species that impart active protection, but also to trap corrosive inonic agents (for example, Cl^-) [13].

2.2. MODA Description of an Inhibitor Performance Simulation Workflow

A representation of the present simulation workflow using MODA [3] is depicted in Fig. 2.

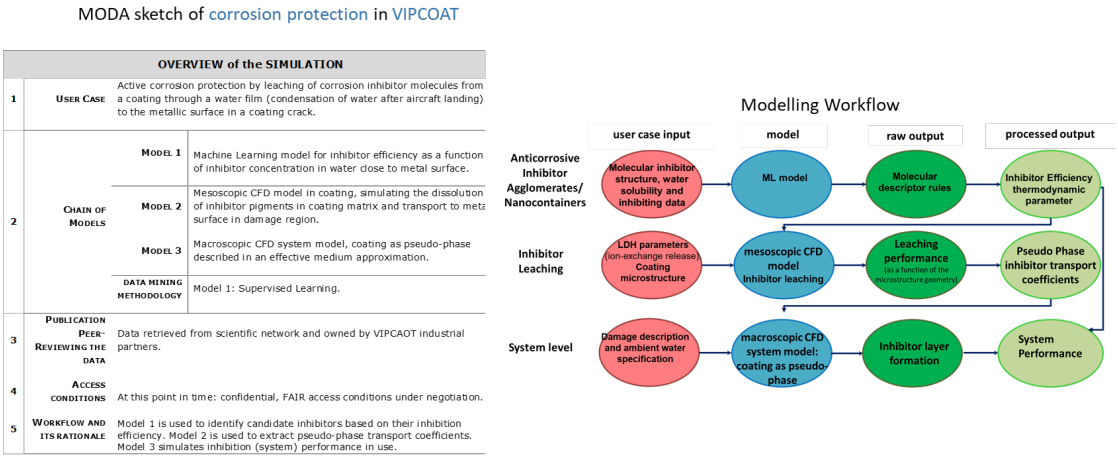

Figure 2: Sketch of the MODA overview section for a simulation workflow of inhibitor selection, inhibitor leaching and inhibitor performance at a system level.

The individual models in the MODA workflow are identified along the following lines of reasoning:

- The coating applied to a metal substrate consists of a polymer matrix enriched with different solid particles, mostly pigments and active inhibitor agglomerates. The release of inhibiting molecules in the polymer matrix is modelled using data based machine learning (ML) as molecular descriptor roles. The data are retrieved from experiments or atomistic modelling using molecular dynamics.
- The leaching within the microstructure is modelled using mesoscopic Computational Fluid Dynamics (CFD). The dissolution of pigments uses the molecular desriptor roles from the previous step as a source term. The microstructure of the coating is fully resolved.
- If a defect occurs and is exposed to humidity, water condenses in the defect area. The leaching model is then coupled to a second CFD model which is responsible for simulating the transport of inhibitor molecules to the substrate surface where they adsorb to form a protective layer. (Fig. 1).

A full MODA/OSMO description of a use case, for each model in the MODA overview section, calls for many more details. In particular, the mathematical equations, their boundary and initial conditions need to be specified, together with a numerical solver used in the MODA workflow. However, the modelling workflow description formalism from MODA [3] is affected by serious

shortcomings and ambiguities [14], particularly in view of an executable workflow deployment. In order to overcome these shortcomings, an alternative approach has been developed, which we outline and apply in the following sections.

2.3. Physical Topology

The physical topology is a graphical method to capture the essentials of a model. It shows the process as a network of extensive-quantity-exchanging capacities. The capacities may be distributed, meaning the intensive properties are a function of the position or lumped if that is not the case. They are also associated with a relative time scale, namely event, dynamic or constant/static. The surroundings of a process are modelled by a set of infinite capacities, which – in analogy to thermodynamics – we call "reservoirs." Only the intensive properties are known for reservoirs, and the total extensive quantities, *i.e.*, mass and energy, are not balanced. It is the intensive properties of the reservoir that make the embedded process change its state. This distinction between extensive and intensive variables and their roles in materials modelling is completely absent from MODA.

In the demonstration case, *cf.* Section 2.1, we use reservoirs (infinitely large with constant intensive properties) to model the electrolyte resources and distributed systems for the water layer on top, the defect, the combination of coating matrix and electrolyte-containing pores and the inhibitor layer. The adsorption of the inhibitor to the substrate is modelled as a point capacity where the adsorption takes place and distributed system for the inhibitor layer (Fig. 3). More details on the graphical modelling language and its relation to MODA can be found in Preisig [6]. The main criticism of MODA here is that there are no control structures in the MODA workflow description leading to an umbiguous mapping to executable workflows.

The inhibitor is stored in a sandwich-type of structure, which captures the inhibitor between layers of support. The support is not exchanging material with its environment, as shown in Fig. 4. Its sole purpose is to contain the inhibitor through a static electrical field, which we do not show in this exhibition. One may view these particles like cookies with a filling of inhibitor. The release of the inhibitor is an ion-exchange process in which the inhibitor ion is replaced by either water or electrolyte components that cause corrosion of the substrate metal with chlorine being the main one. The production of the inhibitor yields agglomerates of inhibitor-loaded cookies, which yield a cumulative interface of the inhibitor particles with the water phase in the pseudo-matrix-water phase (Fig. 3).

2.4. The Workflow as a Cognitive Process

The modelling workflow evolves directly from the topology. As an example, we divided the distributed systems in Fig. 3 into two CFD simulations modelling the behaviour of the pseudophase and the leaching process. Additionally, the leaching process is considered by a molecular simulation of a single LDH cookie.

The physical topology, representing one possible model for the coating problem, can be mapped to a workflow; the mid-level ontology representation of that workflow, using the Physicalistic Interpretation of Modelling and Simulation Interoperability Infrastructure (PIMS-II), is depicted in Fig. 5. Following design choices from the EMMO, the PIMS-II mid-level

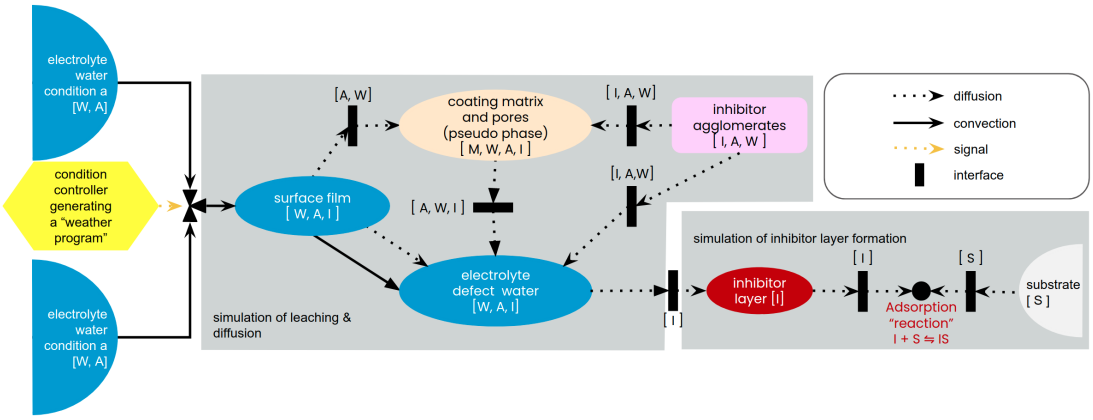

Figure 3: A possible physical topology for the coating process. The "weather" is here simulated by mixing two streams of electrolyte, allowing for changing concentration as well as switching on/off the supply. The coating matrix and the pores are combined to form a "pseudo phase," which implies that properties are normalised by the overall volume. The inhibitor leaches into the pores filled with the watery electrolyte in the matrix, thus the pseudo-phase, and forms an inhibiting layer. The large grey boxes suggest a possible simulation arrangement. Thus, one would simulate the leaching and the diffusion in the crack and the pores together. The building of the inhibitor layers would be solved in a second task.

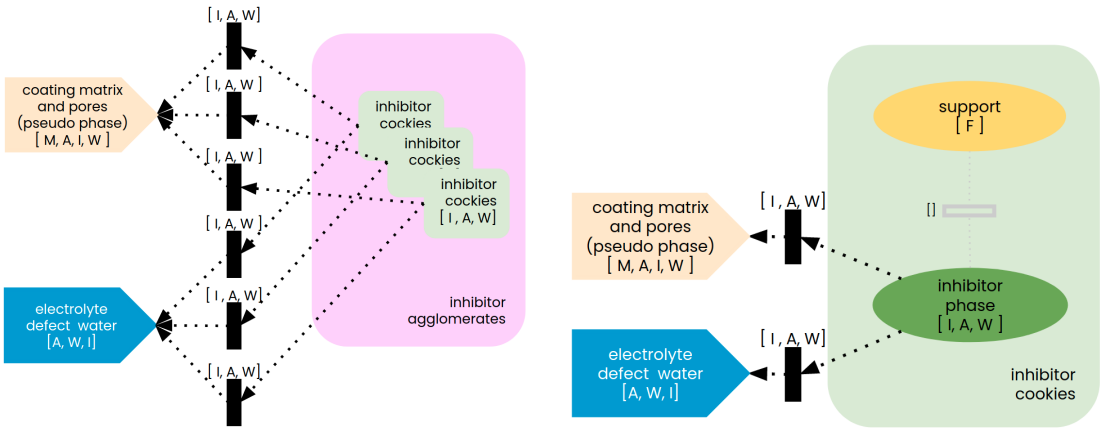

Figure 4: Left: Expansion of the "inhibitor agglomerates" from Fig. 3, showing three representative "cookies" as an example. **Right:** Single cookie, with the support material and the in-between sandwiched ion layer of inhibitor (LDH), active components like chlorine anions, and water.

ontology is fundamentally based on mereotopology [17] and Peircean semiotics [18], which are combined to *mereosemiotics* as a coherent ontological paradigm [19] that is formalized both in OWL2 description logic as well as by a series of axioms in modal first-order logic [15]. In PIMS-II, following Peirce's approach, an elementary cognitive step is a process which is conceptualized as a *triad* and starts from the previously established *representation relation* between a sign (*i.e.*,

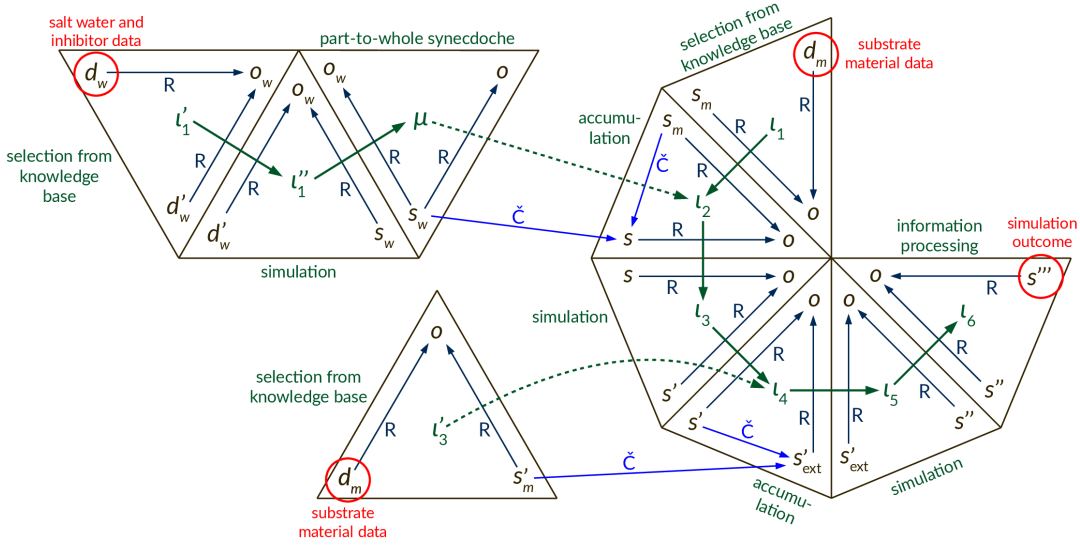

Figure 5: Modelling and simulation workflow represented as a cognitive process, based on Peircean semiotics, using the graphical notation associated with PIMS-II [15, 16]; arrows labelled Č denote the PIMS-II relation isSemioticallyConstitutiveOf, arrows labelled R denote isRepresentamenFor, unlabelled solid arrows denote directlyGrounds, and unlabelled dashed arrows denote directlyPrecedesStep.

a representamen) and an object (*i.e.*, the referent of the representamen); the cognitive step adds a third element to the sign and the object, by which a new representation relation is created. In Fig. 5, cognitive steps are visualized as triangles, where the three triadic elements are situated in the corners; representation relations are denoted by blue arrows from the representamen to the referent, and dependency relations between cognitive steps are denoted by green arrows.

A minimal representation of the workflow is a Petri-like net, which imposes causality: An input/output mapping can only be executed if all inputs are available as indicated in Fig. 6. Therefore, each computation step is equipped with an input gate. The gate functions like a Petri transition: the computation is only started if all inputs are present. In addition, input data must be persistent. Thus, the gate-compute mechanism guarantees causality. Workflows would commonly step in time, which implies that the computational sequences are repeated. Thus computation loops are formed, which must be initialised at the beginning. For this purpose, an additional element, namely an input-selection switch, needs to be introduced, which will change from initial conditions to taking the initial conditions for the next step as the final condition of the previous time step. One may also implement a split, where a signal is passed to more than one gate. An example is the starting point, which sends the starting signal to all three activities.

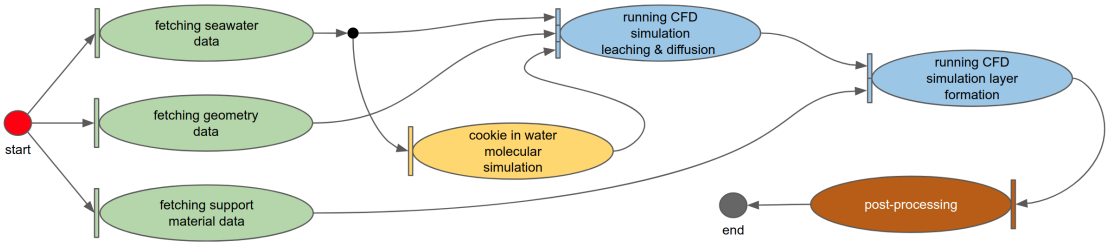

Figure 6: A slightly simplified Petri-like net representation of the workflow leaving out the "weather" simulation section. Each computational step has an import gate. A computation can only be started once all the inputs are available. Data at the input are persistent. A computational step may have several outputs allowing to trigger other computational steps. The workflow is for one time-step only and assumes that the leaching model is computed only once during each step. Adding time-stepping requires a loop, which needs to be initiated by introducing a input switch as an additional control element. The post-processing step includes activities like graphical representation of the result or the computing of performance criteria.

3. MODA to EMMO Crosswalk

3.1. Knowledge Graph Transformation Systems (KGTS)

By a *crosswalk*, we refer to any algorithm, tool, or specification by which instances of one semantic artefact are systematically mapped to instances of another [20]. This includes conventional ontology alignments, obtained as solutions of the ontology matching problem [21], but can go beyond that, as it is indeed necessary for workflow and provenance metadata in computational engineering; *n.b.*, this is not due to any specific complexity of the underlying disciplinary matters. Instead, it is a consequence of requirements by the European Commission, which simultaneously endorses both MODA [3] and the EMMO [1, 2] as metadata standards to be applied to the same domain of knowledge, requiring a crosswalk between them. In the crosswalk from MODA to the EMMO, it is an intermediate step that turns out to be crucial to the quality of the outcome. On the one hand, MODA has an immediate correspondence with OSMO, the domain-ontology version of MODA [14, 22]; on the other hand, the mid-level ontology PIMS-II is designed to be close to the top-level ontology EMMO [15]. However, it is the alignment of OSMO with PIMS-II that it is comparably challenging to implement practically. This reflects a substantial discrepancy between the way in which information is arranged in MODA and the system of relations provided by the EMMO; due to major deviations between the *structure* of knowledge graphs that correspond to each other, applying conventional ontology alignments would lead to a major loss of knowledge. Instead, the present work explores a route based on knowledge graph transformation systems (KGTS), *i.e.*, graph transformation systems [23, 24] applied to knowledge graphs based on RDF triples.

Graph transformation is among the formalisms that are occasionally (though not very frequently) used in semantic-web architectures [26, 27, 28], including work by Mahfoudh et al. [27] on ontology merging. To assess the viability of KGTS for crosswalks between workflow representations in computational molecular engineering, a candidate fragment of a KGTS was

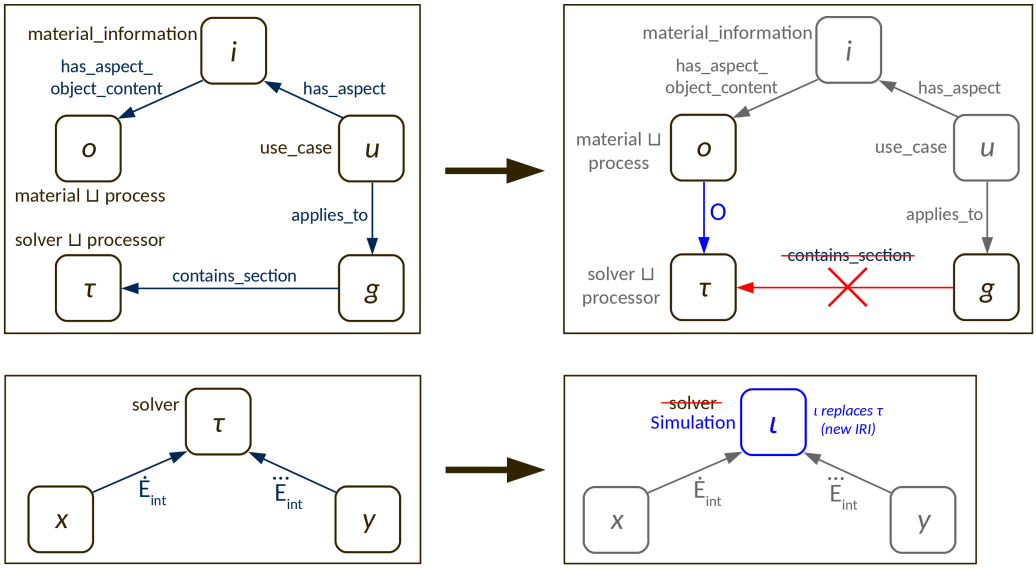

Figure 7: Two rewriting rules selected from a knowledge graph transformation system (KGTS) that aligns OSMO [14, 22, 25] with PIMS-II [15, 16]. Arrows labelled $\dot{\mathsf{E}}_{\mathsf{int}}$ denote the PIMS-II relation isSignInInterpretation, arrows labelled $\ddot{\mathsf{E}}_{\mathsf{int}}$ denote isInterpretantInInterpretation, and the arrow labelled O denotes isReferentIn. The left-hand side specifies the pattern upon which the applicability of the rule is conditional. The right-hand side specifies the graph which with it is to be replaced; grey colour: Vertices and labels unaffected by the transformation step. **Top:** Where the pattern on the left-hand side is detected, the edge labelled contains_section is deleted and an edge labelled O is created. **Bottom:** Where a vertex (individual) instantiating the concept solver has any edges labelled $\dot{\mathsf{E}}_{\mathsf{int}}$ and $\ddot{\mathsf{E}}_{\mathsf{int}}$, representing input and output from a simulation step, it is replaced with a *new* vertex, *i.e.*, an individual with a new IRI is introduced, satisfying all properties and occurring in the same assertions as the old individual; then for the new vertex, the instantiation of solver (from OSMO) is deleted, and an assertion is created by which it instantiates Simulation (from PIMS-II). This replacement is necessary because the solver and the simulation cannot be the same individual, but it is desirable for the new element to retain the role of the old element in several respects; other rewriting rules can be applied subsequently to correct any undesired properties that have been carried over from the source individual.

developed for OSMO as the source ontology and PIMS-II as the target ontology; two selected rewriting rules from this fragment are shown in Fig. 7. In all rules, newly created vertices and edges exclusively instantiate concepts and relations from the target ontology, and *each rule deletes at least one instance of a concept or relation from the source ontology.* This substantially restricts the expressive capacity in comparison with graph grammars in general, which are Turing-complete; *n.b.*, however, that the expressive capacity of conventional ontology alignments is still strictly included, while termination after a linear number of transformation steps is guaranteed. This reduces the problem to $\mathcal{O}(mn)$ instances of the graph isomorphism problem, where n is the size of the source graph and m is the number of rules (usually, a constant). While that problem is not known to be solvable in polynomial time, it has been shown to be quasipolynomial [29]. It is also the use case for which engines of semantic technology software (*e.g.*,

SPARQL end points) are best optimized. Since race conditions can occur between critical pairs of rewriting rules [30], the outcome is not in general uniquely defined; if required, uniqueness of the mapping can be enforced by imposing an order of precedence between applicable rules.

The rules from Fig. 7 illustrate how, particularly by specifying multi-node shape constraints (implementable straightforwardly in SPARQL or SHACL) as a source pattern, the KGTS can retain information that would be lost in a conventional alignment based on immediate conceptual or relational subsumptions: The connection between a simulation (*i.e.*, a semiosis, following Peircean semiotics on which the EMMO is based) and the simulated object is one of the foundational elements for the EMMO, and hence for PIMS-II, which it is crucial to preserve. However, in MODA, and hence in OSMO, it is *not the simulation but the solver* that is immediately represented as a section (corresponding to a page in MODA), and information on the simulated object is *not directly associated with the solver* at all, but with a different section, namely, the use case. Graph-based patterns can take such indirect connections into account.

3.2. Application to the Use-Case Scenario

This is illustrated here for the VIPCOAT use case from Section 2.1, applying the candidate KGTS to the part of the workflow consisting of the CFD simulation of leaching and diffusion and the preceding data-combination processing step. The source graph, shown in Fig. 8 (top), corresponds to an annotation by metadata following OSMO, *e.g.*, as it would be obtained by digitalizing the MODA input provided by a user of a simulation hub or a research data infrastructure. The target graph, obtained by applying a sequence of rewriting rules, is shown at the bottom. The KGTS crosswalk succeeds at retaining the most relevant features from the source graph. This includes information on the dependency (linking) between the two steps of the workflow, the data items communicated from step to step, and the relation between the representing elements and their referent, the simulated leaching and diffusion process. The individuals shared by both knowledge graphs, *i.e.*, nodes that do not undergo deletion or replacement (m, o, s, s', s_m, s_w, and u) during this crosswalk, are highlighted with double-line borders. The capability to replace source individuals with new individuals (as in the second rule from Fig. 7) substantially increases the viability of the crosswalk; *e.g.*, where MODA/OSMO indicates the presence of a solver in a workflow, this almost always means that in the corresponding PIMS-II workflow there should be a Simulation (at EMMO level, a Semiosis). However, an ontology alignment that would subsume one under the other, osmo:solver $\sqsubseteq$ emmo:Semiosis, would be incorrect; the solver and the semiosis are not the same individual, it is only the presence of one element in the source graph that indicates the presence of the other element in the target graph. Such correspondences go beyond what can usually be realized by conventional ontology matching.

3.3. Metadata Curation

While metadata tools can assist in digitalizing and annotating data, it is advisable to include human support in the process. This includes data stewardship [31] and data curation [32], functions that overlap but differ in nuance. Analysing requirements from high-performance computing, Schembera and Durán [7] argue that technical and organizational (or "ethical") tasks interrelate, which is best addressed by advancing the role of a Scientific Data Officer (SDO)

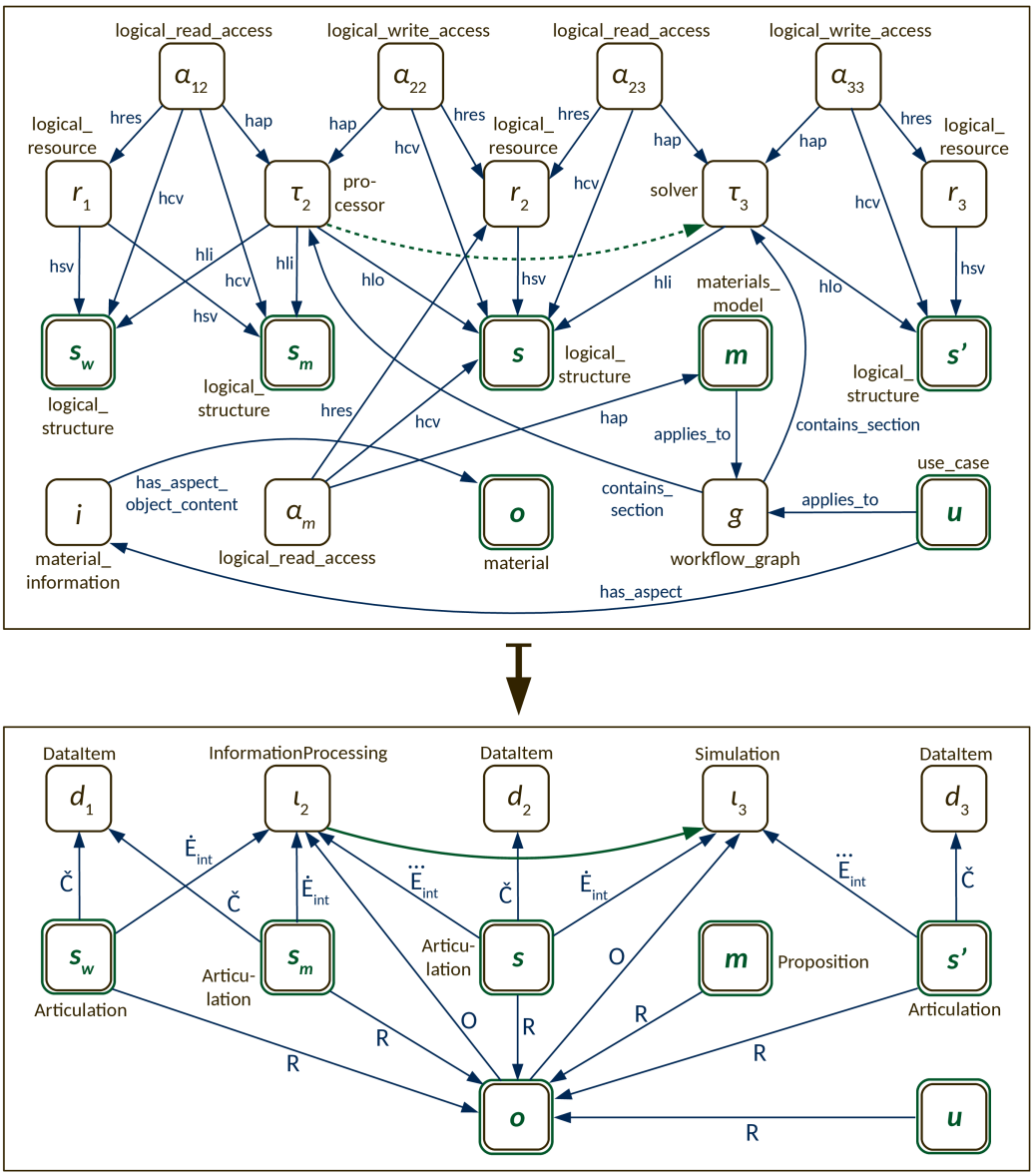

Figure 8: Source and target knowledge graphs from a potential application of a KGTS implementation of the OSMO to PIMS-II crosswalk. **Top:** Source knowledge graph. Abbreviations for relations from OSMO: "hap" for has_access_point, "hcv" for has_carried_variable, "hli" for has_logical_input, "hlo" for has_logical_output, "hres" for has_resource, and "hsv" for has_stored_variable; the unlabelled dashed arrow denotes is_linked_to. **Bottom:** Target annotation in PIMS-II as it would be constructed automatically by an automated crosswalk implementation. Arrows labelled Č denote the PIMS-II relation isSemioticallyConstitutiveOf, arrows labelled $\dot{E}_{int}$ denote isSignInInterpretation, arrows labelled $\ddot{E}_{int}$ denote isInterpretantInInterpretation, arrows labelled O denote isReferentIn, arrows labelled R denote isRepresentamenFor, and the unlabelled solid arrow denotes directlyGrounds.

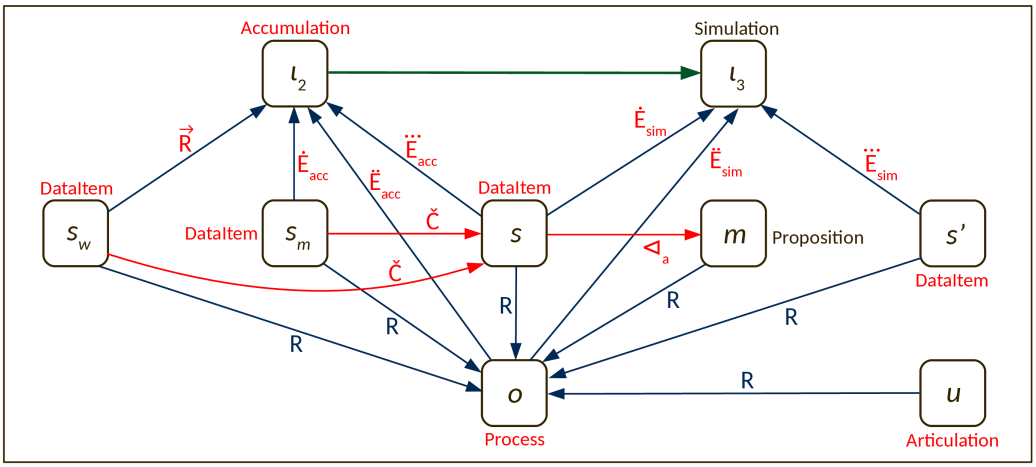

Figure 9: Improved target graph as it might be obtained from an autogenerated target graph as in Fig. 8 (bottom) by the stewardship and curation work of a Scientific Data Officer, *cf.* Schembera and Durán [7]; arrows labelled Č denote the PIMS-II relation isSemioticallyConstitutiveOf, the arrow labelled $\dot{E}_{acc}$ denotes isAccumulationInputIn, the arrow labelled $\ddot{E}_{acc}$ denotes isObjectInAccumulation, the arrow labelled $\dddot{E}_{acc}$ denotes isAccumulationOutputIn, the arrow labelled $\dot{E}_{sim}$ denotes isSimulationInputIn, the arrow labelled $\ddot{E}_{sim}$ denotes isSimulatedIn, the arrow labelled $\dddot{E}_{sim}$ denotes isSimulationOutputIn, arrows labelled R denote isRepresentamenFor, the arrow labelled $\vec{R}$ denotes isIncludedInAccumulation, the arrow labelled $\triangleleft_a$ denotes articulates, and the unlabelled solid arrow denotes directlyGrounds.

as a career. The job description for an SDO encompasses data stewardship and curation, legal and procedural control of good practices, and user management. In the present case, an SDO might reannotate the outcome from the KGTS (Fig. 8, bottom) as illustrated in Fig. 9. Thereby, three kinds of improvements are made: First, instantiations of concepts and relations are made more specific (*e.g.*, from InformationProcessing to Accumulation). Second, the graph structure is simplified by eliminating unnecessary nodes. Third, helpful additional edges are created; here, the model employed in simulation ι_3 occurs in two ways: As the data item s, and as the proposition m. This connection is made explicit by stating that s articulates m (relation $\triangleleft_a$).

4. Conclusion

In this work, we apply an ontology-based toolset to support the *translation* process as specified and recommended by the European Materials Modelling Council (EMMC). Through a minimalistic graphical modelling language with only few construction elements, highly complex model topologies of industrially relevant processes can be modelled; the present work demonstrates this for active protective coating development, applicable to many different surfaces in their respective operational environment. The connection between a MODA-based and an EMMO-based annotation of workflows constitutes a challenge, which the present work resolves by constructing a context-sensitive crosswalk between the two semantic artefacts.

The present KGTS candidate fragment produces acceptable results even without human

intervention. Despite illustrating the potential for automatisation of a translation process and the viability of the suggested approach, it is nonetheless not advised here that complex crosswalks should be deployed in an unsupervised way. Requirements for human oversight apply particularly strongly to the challenge of mapping information content from one metadata schema to another. The difficulty is precisely due to semantic heterogeneity: Conceptual schemes *appear to be incommensurable* unless a crosswalk has already been accepted as valid by the community of its users (which is an organizational task, requiring an agreement) or mappings are approved on a case-by-case basis, requiring an explicit control and affirmation *each time*. It is therefore strictly impossible to substantiate the validity of crosswalks purely by formal verification. If high standards of correctness are to be met (which is certainly not always the case, since approximate annotation is often good enough), human supervision by a translator or an SDO will be unavoidable.

In this work, we demonstrated that it is possible to devise ontology based model construction tools for the specific example of active protective coatings. In the future, we plan to explore more industrially relevant application areas in a similar spirit, most probably using a generalization of the approach used in this work, in order to attain much broader translation capabilities.

Acknowledgments

The co-authors P.K., H.A.P., and N.K. acknowledge funding from the Horizon 2020 research and innovation programme of the EU by grant agreement no. 952903, VIPCOAT, and H.A.P. also acknowledges funding from Horizon 2020 by grant agreement no. 760173, MarketPlace. The co-author M.T.H. acknowledges funding by DFG project no. 441926934, NFDI4Cat, within the NFDI programme of the German Joint Science Conference (GWK). This work was facilitated by activities of Inprodat e.V., Kaiserslautern.

References

[1] J. Francisco Morgado, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, J. Friis, A. de Baas, 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, 2020, p. 3.

[2] E. Ghedini, G. Goldbeck, J. Friis, et al., Elementary Multiperspective Material Ontology (EMMO), 2021. URL: https://github.com/emmorepo/410EMMO.

[3] CEN workshop agreement 17284:2018 E, Materials Modelling: Terminology, Classification and Metadata, 2018.

[4] D. Hristova-Bogaerds, P. Asinari, N. Konchakova, L. Bergamasco, G. Goldbeck, D. Hoeche, O. Swang, G. J. Schmitz, EMMC Translators Guide, 2019. doi:10.5281/zenodo.3552260.

[5] P. Klein, N. Konchakova, D. G. Hristova-Bogaerds, M. Noeske, A. Simperler, G. Goldbeck, D. Höche, Translation in Materials Modelling – Process and Progress, 2021. doi:10.5281/zenodo.4729918.

[6] H. A. Preisig, Ontology-based process modelling – with examples of physical topologies, Processes 9 (2021) 592. doi:10.3390/pr9040592.

[7] B. Schembera, J. M. Durán, Dark data as the new challenge for big data science and the introduction of the scientific data officer, Philosophy & Technology 33 (2019) 93–115. doi:10.1007/s13347-019-00346-x.

[8] P. Visser, Novel totally chrome free corrosion inhibiting coating technology for protection of aluminium alloys, Transactions of the IMF 89 (2011) 291–294. doi:10.1179/174591911X13172025236839.

[9] M. Zheludkevich, J. Tedim, M. Ferreira, "Smart" coatings for active corrosion protection based on multi-functional micro and nanocontainers, Electrochimica Acta 82 (2012) 314–323. doi:https://doi.org/10.1016/j.electacta.2012.04.095.

[10] P. Visser, H. Terryn, J. Mol, On the importance of irreversibility of corrosion inhibitors for active coating protection of AA2024-T3, Corrosion Science 140 (2018) 272–285. doi:10.1016/j.corsci.2018.05.037.

[11] Virtual Open Innovation Platform for Active Protective Coatings Guided by Modelling and Optimization (VIPCOAT, Grant Agreement No 952903) , 2021. URL: https://cordis.europa.eu/project/id/952903/de,https://ms.hereon.de/vipcoat/.

[12] S. K. Poznyak, J. Tedim, L. M. Rodrigues, A. N. Salak, M. L. Zheludkevich, L. F. P. Dick, M. G. S. Ferreira, Novel inorganic host layered double hydroxides intercalated with guest organic inhibitors for anticorrosion applications, Applied Materials & Interfaces 1 (2009) 2353–2362. doi:10.1021/am900495r.

[13] F. Maia, J. Tedim, A. D. Lisenkov, A. N. Salak, M. L. Zheludkevich, M. G. S. Ferreira, Silica nanocontainers for active corrosion protection, Nanoscale 4 (2012) 1287–1298. doi:10.1039/C2NR11536K.

[14] 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, Semantic interoperability and characterization of data provenance in computational molecular engineering, Journal of Chemical & Engineering Data 65 (2020) 1313–1329. doi:10.1021/acs.jced.9b00739.

[15] M. T. Horsch, Mereosemiotics: Parts and signs, in: Proceedings of JOWO 2021 (FOUST 2021), to appear, 2021. Preprint doi:10.5281/zenodo.4849611.

[16] M. T. Horsch, Mereosemiotics: Five scenarios, technical report 2021–B, Inprodat e.V., Kaiserslautern, 2021. doi:10.5281/zenodo.4679522.

[17] N. Asher, L. Vieu, Toward a geometry of common sense: (A semantics and a complete axiomatization of mereotopology, in: Proceedings of IJCAI-95, 1995, pp. 846–852.

[18] C. S. Peirce, Logic as semiotic: The theory of signs, in: J. Buchler (Ed.), Philosophical Writings of Peirce, Dover, New York, 1955, pp. 98–119.

[19] M. T. Horsch, S. Chiacchiera, B. Schembera, M. Seaton, I. Todorov, Semantic interoperability based on the European Materials and Modelling Ontology and its ontological paradigm: Mereosemiotics, in: Proceedings of WCCM-ECCOMAS 2020, Scipedia, 2021. doi:10.23967/wccm-eccomas.2020.297.

[20] T. Habermann, Mapping ISO 19115-1 geographic metadata standards to CodeMeta, PeerJ Computer Science 5 (2019) e174. doi:10.7717/peerj-cs.174.

[21] J. da Silva, F. Araujo Baião, K. Revoredo, J. Euzenat, Semantic interactive ontology matching: Synergistic combination of techniques to improve the set of candidate correspondences, in: P. Shvaiko, J. Euzenat, E. Jiménez Ruiz, M. Cheatham, O. Hassanzadeh (Eds.), Proceedings of OM17, CEUR-WS, Aachen, 2017, pp. 13–24.

[22] M. T. Horsch, D. Toti, S. Chiacchiera, M. A. Seaton, G. Goldbeck, I. T. Todorov, OSMO: Ontology for Simulation, Modelling, and Optimization, in: FOIS 2021 Ontology Showcase (submitted), 2021. Preprint doi:10.5281/zenodo.5084394.

[23] B. König, J. Stückrath, A general framework for well-structured graph transformation systems, in: P. Baldan, D. Gorla (Eds.), Proceedings of CONCUR 2014, Springer, Heidelberg, 2014, pp. 467–481. doi:10.1007/978-3-662-44584-6_32, ISBN 978-3-662-44583-9.

[24] B. König, D. Nolte, J. Padberg, A. Rensink, A tutorial on graph transformation, in: R. Heckel, G. Taentzer (Eds.), Graph Transformation, Specifications, and Nets, Springer, Cham, 2018, pp. 83–104. doi:10.1007/978-3-319-75396-6_5, ISBN 978-3-319-75395-9.

[25] M. T. Horsch, S. Chiacchiera, W. L. Cavalcanti, B. Schembera, Data Technology in Materials Modelling, Springer, Cham, 2021. ISBN 978-3-03068596-6.

[26] H. Q. Yu, Y. Hong, Graph transformation for the semantic web: Queries and inference rules, in: H. Ehrig, R. Heckel, G. Rozenberg, G. Taentzer (Eds.), Proceedings of ICGT 2008, Springer, Heidelberg, 2008, pp. 511–513. doi:10.1007/978-3-540-87405-8_49, ISBN 978-3-540-87404-1.

[27] M. Mahfoudh, L. Thiry, G. Forestier, M. Hassenforder, Algebraic graph transformations for merging ontologies, in: Y. A. Ameur, L. Bellatreche, G. A. Papadopoulos (Eds.), Proceedings of MEDI 2014, Springer, Cham, 2014, pp. 154–168. doi:10.1007/978-3-319-11587-0_16, ISBN 978-3-319-11586-3.

[28] M. N. Roulofs, R. Vos, Automatically inferring technology compatibility with an ontology and graph rewriting rules, Journal of Engineering Design 32 (2021) 90–114. doi:10.1080/09544828.2020.1860202.

[29] L. Babai, Graph isomorphism in quasipolynomial time, in: D. Wichs, Y. Mansour (Eds.), Proceedings of STOC '16, ACM, New York, 2016, pp. 684–697. doi:10.1145/2897518.2897542, ISBN 978-1-4503-4132-5.

[30] G. Grochau Azzi, A. Corradini, L. Ribeiro, On the essence and initiality of conflicts in $\mathcal{M}$-adhesive transformation systems, Journal of Logical and Algebraic Methods in Programming 109 (2019) 100482. doi:10.1016/j.jlamp.2019.100482.

[31] M. J. Cruz, S. Kurapati, Y. Turkyilmaz-van der Velden, The role of data stewardship in software sustainability and reproducibility, in: B. van Werkhoven, A. Mendrik, R. van Nieuwpoort, W. Hazeleger (Eds.), Proceedings of IEEE e-Science 2018, IEEE, Piscataway, 2018, pp. 1–8. doi:10.1109/eScience.2018.00009, ISBN 978-1-5386-9156-4.

[32] A. M. Tammaro, K. K. Matusiak, F. A. Sposito, V. Casarosa, Data curator's roles and responsibilities: An international perspective, Libri 69 (2019) 89–104. doi:10.1515/libri-2018-0090.