CEUR-WS.org/Vol-3647/SemIIM2023_paper_13.pdf

Towards Standardised Product Data Exchange with Information Modelling Framework at Siemens Energy

Maja Miličić Brandt[1,*], Foivos Psarommatis[2], Christophe Simon[3], Arild Waaler[2] and Baifan Zhou[4,2]

[1] *Siemens AG, Germany*
[2] *Department of Informatics, University of Oslo, Norway*
[3] *Siemens Energy, Germany*
[4] *Department of Computer Science, Oslo Metropolitan University, Norway*

Abstract
Business processes in a broad range of industries along the value-chain and product life-cycle suffer from significant challenges pertaining to data quality. These challenges include unknown data validity and tolerance ranges, inconsistency across datasets, data incompleteness, and inaccessibility due to data silos. To address these challenges, we need standardised product data exchange between different business units along the value-chain and product life-cycle. Existing modelling languages such as OWL are expressive, but not sufficiently usable for non-semantic experts, while typically these experts (e.g., engineers) possess the essential knowledge for creating the data models. To this end, we develop an approach named as Information Modelling Framework (IMF) that aims at user-friendly modelling for engineers. This paper presents our ongoing research of the IMF approach and exemplify its usage with a real business case at Siemens Energy.

Keywords
information modelling, standardised data exchange, system engineering

1. Introduction

Motivation. In a broad range of industries, including manufacturing, process, etc. the business processes along the value-chain and product life-cycle suffer from significant challenges pertaining to data quality [1, 2]. These challenges include pervasive issues such as unknown data validity and tolerance ranges, inconsistency across datasets, data incompleteness, and inaccessibility due to data silos [3]. The data quality issues have far-reaching implications, affecting critical aspects of operations, decision-making, and efficiency [4]. There is a pressing need to address these data quality concerns by establishing standardised protocols for product data exchange [3]. At Siemens and Siemens Energy, we endeavour to contribute to the resolution of these challenges by exploring and proposing innovative solutions for semantic

---

SemIIM'23: 2nd International Workshop on Semantic Industrial Information Modelling, 7th November 2023, Athens, Greece, co-located with 22nd International Semantic Web Conference (ISWC 2023)

[*] The authors are ordered alphabetically by the last names.

maja.milicicbrandt@siemens.com (M. Miličić Brandt); foivosp@ifi.uio.no (F. Psarommatis); christophe.jt.simon@siemens-energy.com (C. Simon); arild@ifi.uio.no (A. Waaler); baifanz@ifi.uio.no (B. Zhou)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

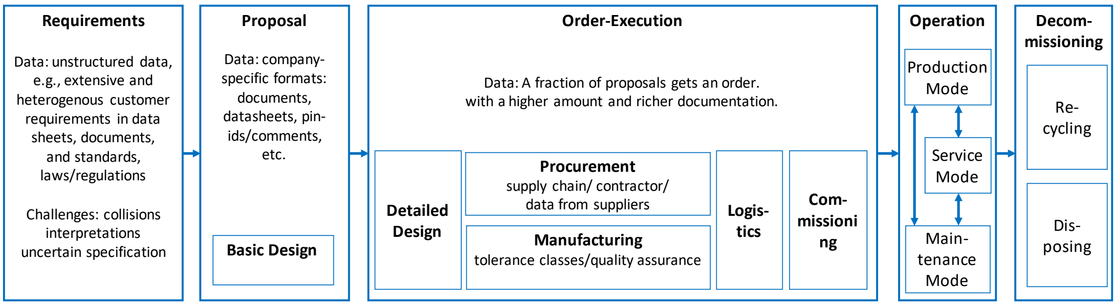

Figure 1: Sketch of activities carried out by different business units along the value-chain and product life-cycle. Note that the figure is a simplified overview.

industrial information modelling, ultimately aiming to enhance data quality, accessibility, and interoperability within Siemens Energy and similar industrial contexts [5].

Challenges. In particular, we consider the following challenges:

- *C1, Consolidation challenge to address semantic ambiguity and data transformation.* The exchange of technical information frequently occurs without standardised semantics and consistent data structures. This process often relies on unstructured documents or datasheets, both in inter-company transactions involving clients and suppliers and within internal teams. The result is a multitude of different data formats that require unification for effective communication.
- *C2, Provenance challenge that results from loss of information along the extended value-chain.* Across the extended value-chain (Fig. 1), valuable technical information can become fragmented and scattered. Multiple copies of the same data fields proliferate, and numerous intermediate steps introduce complexity and potential for data loss. This lack of transparency and data quality hinders operational efficiency and effectiveness.
- *C3, Automation challenge in ensuring data integrity.* Current practices often necessitate manual verification of requirements, a process prone to errors. This manual approach not only consumes valuable resources but also raises concerns about data integrity and accuracy. An automated solution is needed to enhance the reliability and consistency of requirement verification processes.

Requirements. Our goal is to develop an approach to achieve standardised product data exchange between the different units in the business process and value-chain. Based on the industrial/business needs, the requirements include the following points:

- *R1, Standardised/shared*: the approach should have standardised vocabulary and information structure with a good coverage and a process of extending it;
- *R2, Machine-processable*: the approach should have automatable information verification mechanism to ensure that the info satisfies a set of specifications;
- *R3, Precise*: the approach should use unambiguous/precise languages, with minimal redundancy, be clear cross references and well-organised;
- *R4, User-friendly*: the approach should be easy to use for engineers (non-semantic experts);
- *R5, Good interface/modularised*: to use the information modules, users only need to understand the modules, and do not need to understand the whole content;

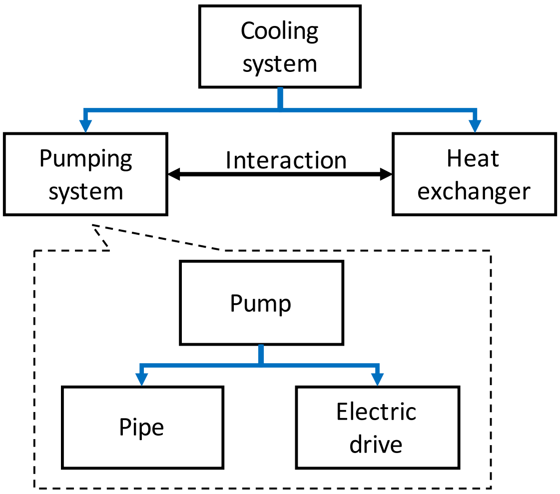

Figure 2: System thinking: breakdown, granularity, and interaction.

- *R6, Flexible/versatile*: the approach should accommodate a broad range of needs/interests;
- *R7, Provenance traceable*: the approach should be table to track where the information comes from, to have clear provenance along the information propagation chain.

Existing Approaches. The existing approaches can be categorised largely into two groups. One group is intended for users who are experts in information modelling, such as OWL [6], SHACL [7], etc. These approaches are standardised, machine-processable, precise, and flexible, but they require extensive training before the users can start to be productive (not meeting R4). The other group includes tools intended for engineers, such as COMOS [8], but more work needs to be done in reaching agreement between organisations and standards (R1), and they are by large not machine-processable (R2) and sufficiently precise (R3).

This paper presents our ongoing research on standardisation of product data exchange between different business units along the value-chain and life-cycle. We rely on the approach of Information Modelling Framework (IMF), which is being developed as an engineering friendly method to model engineered systems and assets. We aim at using IMF models as the standardised data exchange format, and present a gearbox use case at Siemens Energy as preliminary results to exemplify our approach. This new approach has the rigour of logic-based formal languages while being user-friendly enough for engineers, thus bridging the gap between the existing approaches.

2. Approach: Information Modelling Framework

Approach Guidelines. We rely on the approach of Information Modelling Framework (IMF) [3]. IMF is being developed as an engineering friendly method to model engineered systems and assets. IMF aims to bridge the gap between ontologies and industrial system data by:

- Implementing ISO/IEC 81346-1 in RDF [9];
- Explicitly representing Aspects (function, product, location) of the asset along its life-cycle [9];

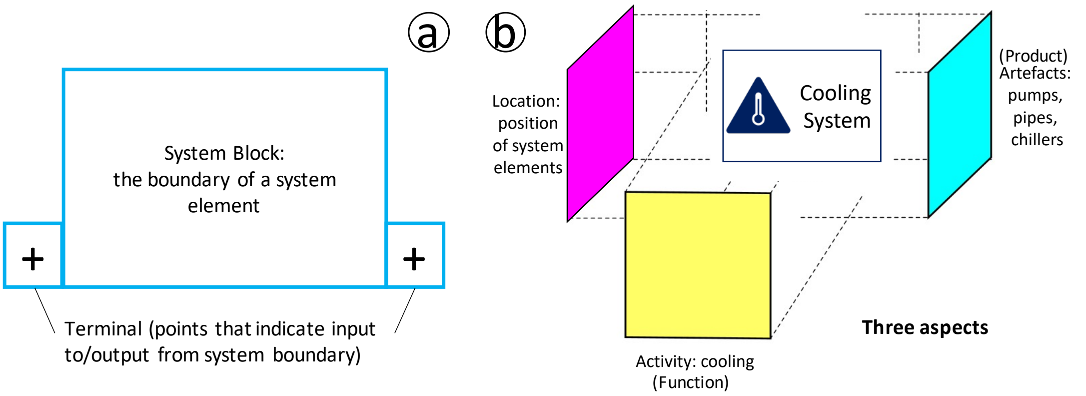

Figure 3: (a) The two graphical elements in IMF: *System Block* (or *Block* for short) and *Terminal*. (b) Three *Aspects* in IMF modelling.

- Using a limited vocabulary to express engineered systems;
- Alignment with Industrial Data Ontology (former ISO 15926-14) [10].

System Thinking. To model engineered systems and assets, we first need some basic concepts of system thinking. For simplicity, we brief these concepts in informal descriptions and take the example of a cooling system (Fig. 2).

- *System and system elements.* A *system* is a group of system elements to achieve a purpose, where a system element is a system by itself. For example, a cooling system achieves the purpose of the cooling function, that is to reduce the temperature of e.g., an physical object.
- *Breakdown and granularity.* A system can be broken down to system elements, and this *breakdown* can be continued to a *granularity* that is needed. The cooling system can be broken down input a pumping system and a heat exchanger, where the pumping system can be further broken down into pumps, which can be broken down into pipes and electric drives. This breakdown goes to a granularity that is needed.
- *Interaction.* The system elements can *interact* with other system elements. For example, the pumping system interact with the heat exchanger via pipes and liquid in the pipes.

IMF Modelling Basics. With the system thinking as the background knowledge, we shortly introduce some IMF modelling principles and refer the readers to the complete manual for details [11]. To model an engineered system in IMF, is to describe the system with connected graphical elements in three Aspects. The syntax is introduced informally as below: [1]

- *Graphical elements.* Aiming a simplistic modelling language for non-semantic experts, especially engineers, IMF provides two language elements, *system block* and *terminal* (there is other language element *interface point*, which is not in focus of in the paper). These language elements are also called graphical elements, because they have graphical representation as shown in Fig. 3a.

---

[1] This paper aims at a minimal description of the IMF modelling principles that is needed for understanding the gearbox case. We thus significantly simplified the IMF modelling principles. The simplified principles still align with the full-extent principles.

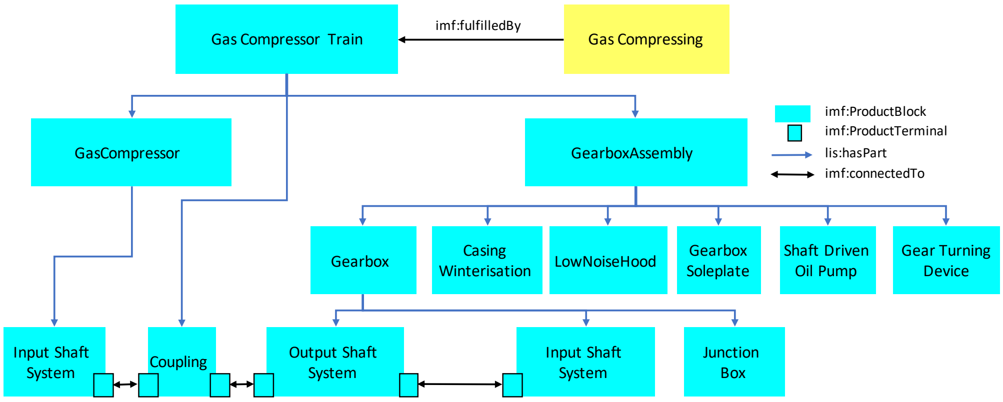

Figure 4: The Product Aspect (excerpt) of the IMF model for the gearbox case

- *Three Aspects.* As in ISO/IEC 81346-1 in RDF [9], a system can be viewed from three Aspects: function, product, and location (Fig. 3b). These are three perspectives of viewing a system:
  - the *Function Aspect* of a system is viewing a system from the perspective of activity, because the function of a system is the potential of the system to realise an activity (the cooling activity in example);
  - the *Product Aspect* of a system is viewing a system from the perspective of artefact, because a system consists of a set of physical artefacts (such as pumps, pipes, chillers);
  - the *Location Aspect* of a system is viewing a system from the perspective of location, which is the (geographical) positions and spatial extension of system and system elements.
- *Relations.* The graphical elements are connected via a set of relations. Here we brief common relations and their domains and ranges that are needed for this paper:
  - `lis:hasPart`. This relation connects (has domain and range of) *Blocks* of the same *Aspect*, or *Terminals* of the same *Aspect*. The relation is used to represent the breakdown of a system. The relation is from Industrial Data Ontology (IDO for short, former ISO 15926-14) [10] (lis is the namespace of IDO).
  - `imf:connectedTo`. This relation has both domain and range of *Terminals* of the same *Aspect*. The relation is used to represent the interaction of a system.
  - `imf:fulfilledBy`. This relation has domain of *Function Block* and range of *Product Block*, or domain of *Function Terminal* and range of *Product Terminal*.

Shared Semantics. IMF aims at semantics shared among the stakeholders. These include three components:

- *Semantics of information models*: Semantics of objects, relations, and attributes is defined in shared IMF models and ontologies, between the stakeholders, such as clients, Siemens Energy, suppliers, etc.
- *Semantics of attributes*: These also include standardised specifications, such as attribute qualifiers, description of a product life-cycle and milestones, solution documentation
- *Semantics of data validation*: Reasoning and data validation are done using OWL reasoners and SHACL engines.

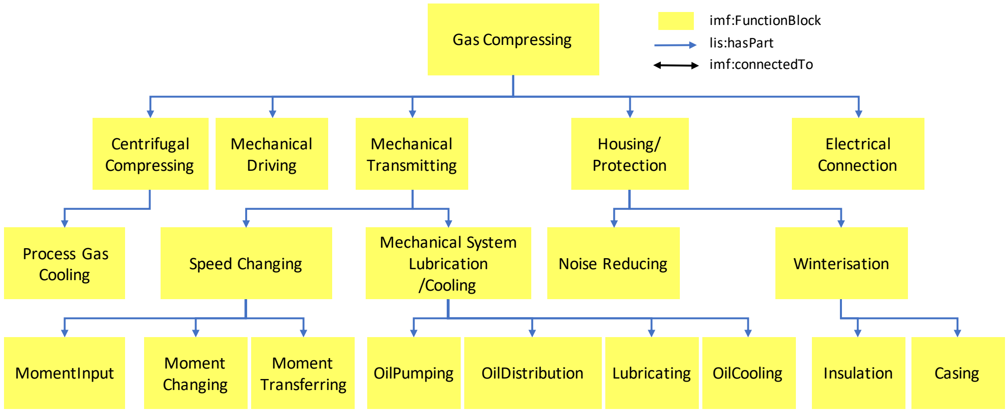

Figure 5: The Function Aspect (excerpt) of the IMF model for the gearbox case

3. Gearbox Use Case

A gearbox is a common component of a gas compressor train that is interconnected with many other components. The gearboxes are normally supplied by external suppliers and incorporated into solutions provided by Siemens Energy. This section presents the modelling of a gearbox in the gas compressor with IMF, aiming at standardised data exchange with the IMF models. We discuss the gearbox as a system element in a gas compressor train. The IMF models of a gearbox includes *Function Aspect*, *Product Aspect*, and *Location Aspect*.

Product Aspect. To ease the understanding, we start with the *Product Aspect*, which describes the breakdown and interaction of the physical artefacts of the system (Fig. 4).

- *Gas Compressor Train* Block. It is the root node of the IMF model in the *Product Aspect*. A gas compressor train is often referred to as a compressor package or gas compression package. It is a complex mechanical system used in various industries to compress natural gas or other gases. The gas compressor train is the overall artefact that fulfils the function of compressing the incoming gas. The *Gas Compressor Train* Block has many parts, such as the *Gas Compressor* Block and the *Gear Box Assembly* Block.
- *Gas Compressor* Block. The gas compressor is the central component of the gas compressor train that fulfils the primary function of compressing the incoming gas.
- *Gearbox Assembly* Block. The gearbox assembly is a critical subsystem that efficiently transmits power from the driver, typically an electric motor or a gas turbine, to the compressor and other auxiliary components. This assembly ensures that the compressor operates at the correct speed and torque, optimising the compression process. *Gearbox Assembly* Block has further six parts. For conciseness, we omit these details.
- *Coupling* Block and connections between the terminals. In Fig. 4, we show the *Terminals* and the `imf:connectedTo` between the *Terminals* partially. These connections indicate that the input shaft system of the gas compressor is connected to the coupling via some media (here force), and the coupling is connected to the output shaft system of the gearbox via media.

Function Aspect. The *Gas Compression* Function Block is fulfilled by the *Gas Compressor Train*

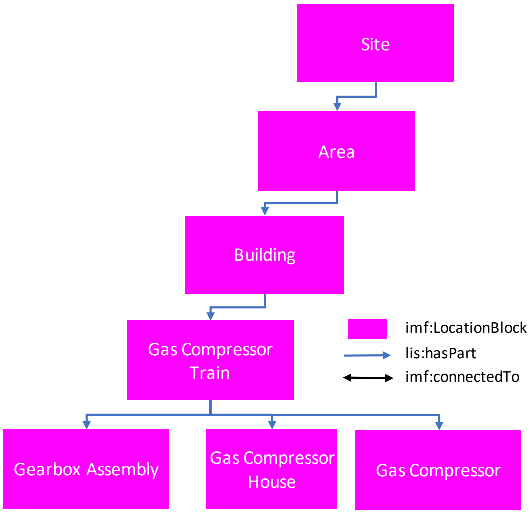

Figure 6: The Location Aspect (excerpt) of the IMF model for the gearbox case.

Product Block. We illustrate the IMF model for the system from the function aspect in Fig. 5.

- *Gas Compression* Block. It is the root node of the IMF model in the *Function Aspect*, which describes the primary and overarching function of a gas compressor, that is to compress gas. This compression process reduces the volume of the gas while increasing its pressure. This function can be achieved if all of its sub-functions are achieved.
- *Centrifugal Compressing* Block is one of the sub-functions, which describes the (potential of) the activity that the compressor uses the mechanical energy to increase the gas pressure.
- *Mechanical Driving* Block describes the (potential of) the activity to provide the mechanical energy.
- *Housing/Protection* Block describes the (potential of) the activity to provide a sturdy and protective casing or enclosure for the compressor's internal components. It is an activity because it spans over time.
- *Electrical Connection* Block describes the (potential of) the activity to establish the necessary wiring and connections for the control system to interact with the compressor.
- These sub-functions can be achieved if their sub-functions are achieved. For conciseness, we omit further details.

Location Aspect. The Location Aspect is the description of installation positions (intended or actual) at the customer site and the product spatial extension.

- The first three blocks always exist for all engineered systems. *Site* Block describes the geographical position and spacial extension of the overall site where the gas compressor operates. *Area* Block describes a part of the site, and *Building* Block describes the building, which is a part of the area.
- The following blocks describe the gas compress train from the location aspect. The *Gas Compressor Train* Location Block has three parts, which is the *Gearbox Assembly* Block, *Gas Compressor House* Block, and the *Gas Compressor* Block. These blocks contain attributes of

the position and spacial extension of each physical artefacts.

Summary. The three figures (Fig. 4 - Fig. 6) are to exemplify how to describe engineered systems with the IMF models. We could see that despite the very limited allowed vocabulary (graphical elements, aspects, and relations), very complex engineering systems from real business case can be described. In addition, because of the limited vocabulary, representations of such systems in IMF models are easy and intuitive for engineers to understand, and close to the way they think. After having the blocks and terminals, the attributes with specified values and ranges are attached to the blocks and terminals, organising most of the essential information of engineering systems.

4. Conclusion and Outlook

Conclusion. With good data quality and standardised data models, we could reduce non-conformance cost, improve productivity (less time spent on search and confirmation) in projects, increase application and data re-use in IT and administration, shorten cycle time of projects, and improve reaction times to third parties. IMF aims at such standardised data models, and the "lingua franca" between organisations, business units along the product life-cycle and value-chain. This work is a step towards a grand goal, and provides preliminary guidelines and experience for future investigation and industrial practice.

Limitations. Looking at the challenges C1-C3, we can see that C1 is partially addressed by the presented IMF approach, while C2-C3 are not touched. Regarding the requirements, we can see that R1, R3, R4, R5, R6 are partially fulfilled by the IMF approach, while R2 and R7 require future research. C3 and R2 require translating the IMF models into some well-defined languages, such as OWL and SHACL. C2 and R7 require more research on tracking the information drift along the product life-cycle and value-chain, especially tracking the attributes changing, because preserving information across the attribute life-cycle is essential in quality assurance.

Outlook. As future work, we see the following directions:

- *semantic interpretation* that translate the IMF models to some well-defined language such as OWL and SHCAL;
- *semantic verification* that automatically validates the data integrity;
- *software engineering* that provides different pieces of tooling and end-to-end tooling for IMF modelling, including Mimir, Tyle (by Equinor) as starting point as user interface, and Domain Ontology Editor (by Siemens Energy) for ontology administration;
- *alignment with various standards*, we need shared ontologies mapped to various standards, such as CFIHOS, AAS, etc.
- An *ecosystem* of industrial information modelling is being built across many organisations.

Contributions by Authors

Storyline: Maja. Motivation, challenges and requirements: Christophe, Maja, Arild, Baifan and Foivos. IMF approach: Arild, Baifan, Foivos, Maja. Gearbox case: Christophe. Figures: Baifan, Maja, Christophe, Foivos. Writing: Baifan, Maja, Christophe, Foivos, Arild.

Acknowledgments

This publication is supported by the European Commission H2020 with the project number 958371 (OntoCommons) and by the Norwegian Research Council with project number 237898 (SFI SIRIUS), 308817 (DigiWell).

References

[1] K. Grevenitis, F. Psarommatis, A. Reina, W. Xu, I. Tourkogiorgis, J. Milenkovic, J. Cassina, D. Kiritsis, A hybrid framework for industrial data storage and exploitation, Procedia CIRP 81 (2019) 892–897.

[2] F. Ameri, D. Sormaz, F. Psarommatis, D. Kiritsis, Industrial ontologies for interoperability in agile and resilient manufacturing, International Journal of Production Research 60 (2022) 420–441.

[3] D. B. Cameron, A. Waaler, E. Fjøsna, M. Hole, F. Psarommatis, A semantic systems engineering framework for zero-defect engineering and operations in the continuous process industries, Frontiers in Manufacturing Technology 2 (2022) 945717.

[4] F. Psarommatis, D. Kiritsis, A hybrid decision support system for automating decision making in the event of defects in the era of zero defect manufacturing, Journal of Industrial Information Integration 26 (2022) 100263.

[5] S. Cho, G. May, D. Kiritsis, A semantic-driven approach for industry 4.0, in: 2019 15th international conference on distributed computing in sensor systems (DCOSS), IEEE, 2019, pp. 347–354.

[6] OWL 2 web ontology language document overview (second edition), W3C Recommendation (2012).

[7] H. Knublauch, D. Kontokostas, Shapes constraint language (SHACL), W3C Recommendation (2017).

[8] Siemens, COMOS software solutions, 2023. URL: https://www.siemens.com/global/en/products/automation/industry-software/plant-engineering-software-comos/portfolio.html, accessed 31 Aug. 2023.

[9] ISO/IEC 81346-1, 81346-1: Industrial systems, installations and equipment and industrial products - Structuring principles and reference designations - Part 1: Basic rules, Standard, International Organization for Standardization/International Electrotechnical Commission, 2022.

[10] IDO, Industrial Data Ontology - New Work Item Proposal under ISO TC 184/SC 4, Standard Draft, PCA, 2023.

[11] E. Fjøsna, T. Saltvedt, A. Waaler, M. Knædal, V. Koppergård, L. Hella, M. G. Skjæveland, R. Mehmandarov, M. Fekete, B. Zhou, Information Modelling Framework Manual – Draft Version 2.1, Manual, 2023. URL: https://www.imfid.org/.