Use of Ontologies to Structure and Manage Digital Technical Data of Industrial Assets: First Steps Towards an Ecology of Knowledge in Multi-Energies Industry

Jean-Charles Leclerc[1], Gregory Tetard[1], Yves Keraron[2], Claude Fauconnet[3]

[1] TotalEnergies, Tour Coupole, 2 place Jean Millier, 92400 Courbevoie, France
[2] ISADEUS, 20 avenue des muses, 44470 Carquefou, France
[3] SousLeSens, claude.fauconnet@gmail.com

Abstract

The use of semantic web technologies and of ontologies is raising growing interest in industry.
Industry is facing during these last decades, with the rapid development and use of a set of applications for specific needs, new issues due to a lack of interoperability, which conducts to growing costs. When digitization of industry is accelerating, ontologies and digital data standards based on models pave the way to solve these issues by improving interoperability between applications and between actors of the value chain.

Recent European H2020 projects and clusters, Joint Industrial Projects in industry and advanced standards of Standardization Development Organizations converge towards new ways of modeling, structuring, and managing digital technical data across applications using ontologies thanks to advanced conceptual, methodological, and semantic technologies.
In this paper, we will provide preliminary results for structuring transverse information knowledge and data in a proper manner using industry standards and semantic framework.

.

Keywords

Industrial ontologies, formal ontology, domain ontology, data interoperability, semantic information modelling, web technological standards.

1. Introduction

Industry is facing growing costs of lack of interoperability and is aware that the use of standards is key for the success of its digital transition.

Efforts have been made, standards and frameworks have been developed for a better integration of digital technical data in process industries with the ISO 15926 series [1] edited by the TC 184 SC 4 of ISO and the ISO/IEC 81346 series [2] of ISO TC 10 and initially published by IEC.

A new ISO/CD TR 15926-14, Data model adapted for OWL2 Direct Semantics [3], is under development and experienced in the joint industrial project READI [4] for a greater alignment with the needs of the whole lifecycle of an asset and with top level ontologies as BFO [5], DOLCE [6] or Tupper [7], already standards or in a process of standardization in the ISO/IEC JTC 1 SC 32.

TotalEnergies has developed its own methodology and an open-source set of tools in a semantic framework we will present in this paper.

FOMI 2022: 12th International Workshop on Formal Ontologies meet Industry, September 12-15, 2022, Tarbes, France
*Corresponding author
Email: jean-charles.leclerc@totalenergies.com (Jean-Charles Leclerc)

2. Challenges of the Management of the Digital Technical Data in Process Industry and the Adopted Approach

Currently, industry faces serious challenges to properly manage its digital technical data, among them: Digital technical are exchanged between the actors without clear specifications, protocols and formats for an efficient reuse of these digital technical data; data are edited in proprietary tools with lock-in formats, which limits the reuse of data and their long-term archiving; most of the standards are not machine readable, which makes it difficult to deploy a standard based digital strategy.

The approach to manage, integrate, and interoperate digital technical data is based on 3 pillars:

- The ISO/CD TR 15926-14, with other standards, is the first pillar to support the integration of the digital technical data. It supports the mapping of the models used for legacy data and applications on the concepts and relationships to be used to perform, link, prescribe and describe according to different aspects the assets.
- The ISO/IEC 81346 is the second pillar of the approach to manage digital technical data. It is a framework to organize and manage digital technical information. A key concept of ISO/IEC 81346 is the concept of aspect, and 3 main aspects are considered: Function, Product and Location. The Function aspect is developed in a performative structure of the asset; at each level of the structure, the functions are allocated to a physical object of the product structure, object which is selected to realize the requirements based on the properties to fulfill the function. The structures according to different aspects support a systems' engineering methodology.
- The development of Reference Data Domains knowledges graph of company Standards where *Business Objects* build constitute the third pillar to interoperate. A rigorous formalization of sets of objects and their relationships as used in legacy applications or in applications controlled by a shared ontology, open the door to the composition of these objects to provide the users with the services they need in their activities.

3. Methodology to Develop the Semantic Framework

The evolution of industrial practices to manage data and knowledge in the context of industry 4.0 and Big Data faces the challenge of integration data together to produce actionable knowledge. Problems are mostly linked both to the heterogeneity and partiality of data models and to the huge difficulties inherent to natural language: polysemy of terms alone, large ambiguity contained by the predicates constituting the speech, reduced in human conversation but hardly processable by machines.

On the other hand, increasing demand of enterprises to integrate and interoperate data is limited by the rigidity and the "subjectivity" of the data models. Interconnecting data in this context leads to complexity and costs growing exponentially with the number of sources to integrate.

TotalEnergies Semantic Framework provides an approach and tools[2] based on a rigorous usage of the semantic web technologies and formal ontologies together, following ISO-15926-Part14 principles. It consists into the elaboration of several domain ontologies (domain knowledge graph models) that semantizes the data related to the domain business objects. The underlying process (figure 1) starts with the mapping of business object terms to the classes of the formal abstract ontology and their linking with properties respecting the formal ontology axioms [3].

[2] SousLeSensVocables : C.Fauconnet license MIT : https://github.com/souslesens/souslesensVocables

[3] For instance, a physical object can only be linked to a location by the property « located in » or a sub-property of it

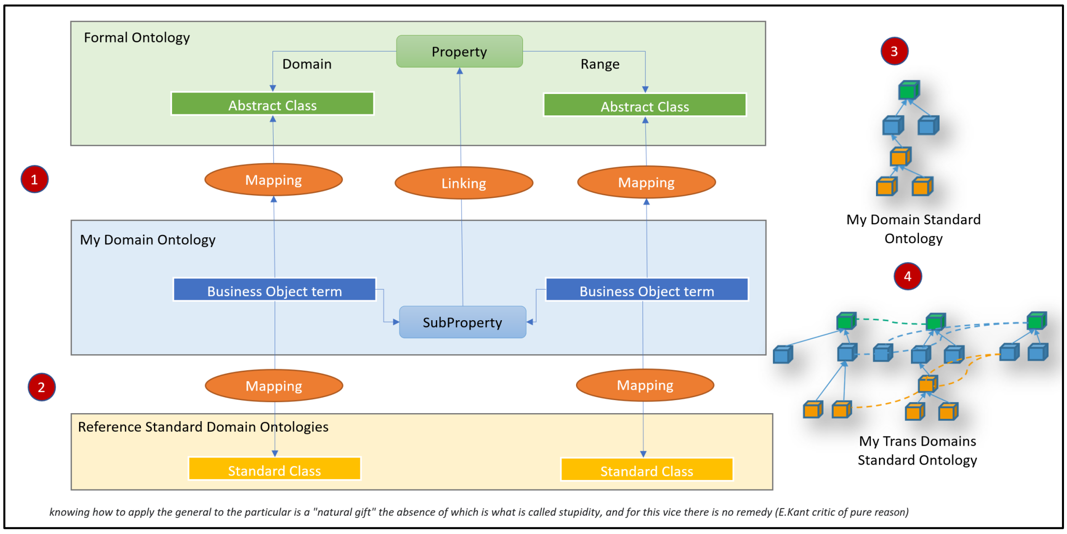

Fig 1: Business Objects mapping methodological modeling process for standardization.

This process appears to be efficient for several reasons: It can be performed by business experts themselves (as a matter of fact they like it!); it disambiguates the concepts radically and name them precisely; the weaving of semantic predicates between classes respecting the axioms of the formal ontology and inheritance mechanisms brings a lot of value, also for future applications; the generated graph is "mostly objective" and pertinent because it links the concepts used by a specific domain of knowledge to universal abstractions that can be shared and agreed by people, building cross domains bridges.
Standardized Knowledge Graphs supported by formal ontology appear key for the elaboration of a strategy aiming at the integration of companies' data and knowledge.

The attachment of domain class to the formal ontology leads to the question of how to organize classes and hierarchies. ISO/IEC 81346-1 makes the proper distinction with two types of hierarchies: the classification hierarchies based on the class-sub-class relationship; the composition hierarchies based on the whole-part relationship. After building composition hierarchies according to various aspects, designating the objects, linking them to classification hierarchies, it is possible to link the nodes of the hierarchies with various relationships. The management of this digital technical information ecosystem is controlled by the axioms and the constraints of the formal ontology based on ISO 15926-14.

4. A Suite of Semantic Tools and Their Use

We already introduced our TotalEnergies Semantic Framework tools and methodology at I-ESA 2022 conference in a paper in publication [8]. This methodology towards formalization of knowledge of the company: Business references as data domains standardized objects, called *business-objects* are managed through classification and composition hierarchies.

The amount of duplicated data and information is generally under-evaluated; one obstacle in this bottom-up approach was the disinterest of the business with the data they are producing/consuming because of the necessary domain knowledge to manage data architecture of Conceptual, Logical & Physical data models, wrongly considered as an IT specialist matter only. The recent reconsiderations of the data and the associated new potential of growth of data driven companies lead to reconsider previous practices and weaknesses around data management.

The figure 2 below shows an example of disambiguation of concepts from a partial sketch for an Exploration Data Domain's ontology based on ISO15926-14 formal ontology.

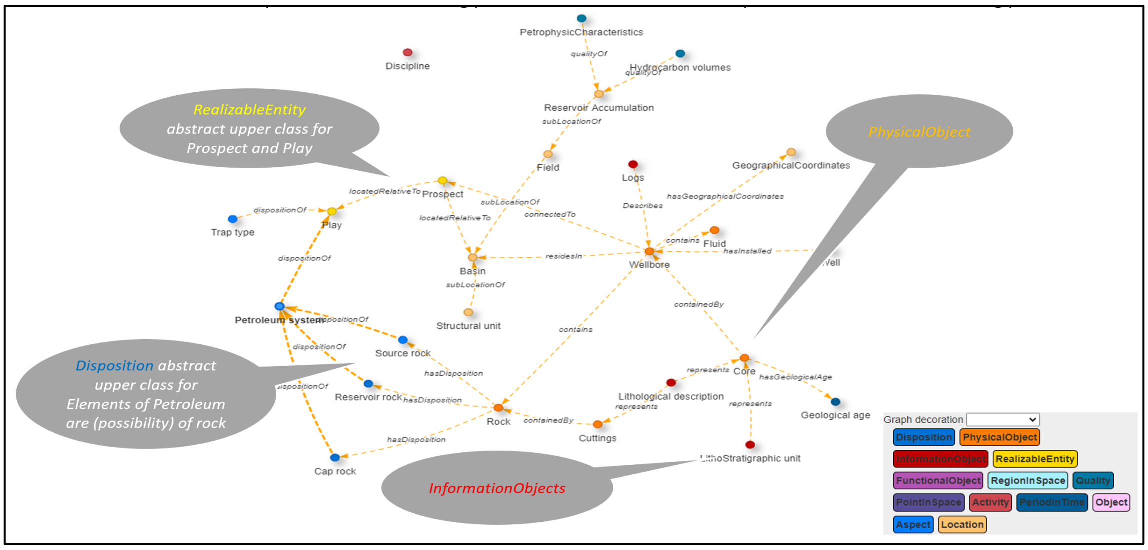

Fig 2 : Sample of Geosciences Data Domain linked business objects related to ISO15926-14.

5. Outlook and Conclusion

Industry is a domain where a considerable amount of knowledge is distributed in various tools and in the mind of the supply chain over time. The extensive use of software tools to edit and collect data face serious problems of lack of interoperability. Common principles, methodologies, and tools to manage technical information and to align the meaning of the concepts used in models and standards is progressing thanks to the works carried out in professional and standards development organizations, which will be a key success factor towards an ecology of knowledge in a multi-energies company.

First results are encouraging, and a lot of work remains to be done.

The next steps will consist of:

- extensive testing of the capabilities of ISO 15926-14 to model and link various kinds of business objects, commons of TotalEnergies, critical to make accessible the knowledge all over the company,
- combination of different techniques as natural process language and ontologies to make accessible the knowledge embedded in the documents.

6. References

[1] ISO 15926:2004, Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 1: Overview and fundamental principles

[2] ISO/IEC 81346, Industrial systems, installations and equipment and industrial products — Structuring principles and reference designations — Part 1: Basic rules

[3] ISO/CD TR 15926-14, under development, Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 14: Data model adapted for OWL2 Direct Semantics.

[4] READI, REquirements And Digital Information, Shaping The Future Of Digital Requirements And Information Flow In The Oil And Gas Value Chain: https://readi-jip.org

[5] BFO, ISO/IEC 21838-2:2021, Information technology - Top-level ontologies (TLO) - Part 2: Basic Formal Ontology (BFO)

[6] DOLCE, ISO/IEC DIS 21838-3, Information technology - Top-level ontologies (TLO) - Part 3: Descriptive ontology for linguistic and cognitive engineering (DOLCE)

[7] TUPPER, ISO/IEC DIS 21838-4, Information technology - Top-level ontologies (TLO) - Part 4: Tupper

[8] New ways of using standards for semantic interoperability towards integration of data and models in industry, Yves Keraron, Jean-Charles Leclerc, Claude Fauconnet, Nicolas Chauvat, Martin Zelm, I-ESA 2022, Valence, Spain.