Towards Standardization of Ontologies in Research and Industry

Frank-Walter Jaekel
Institute for Production Systems and Design Technology,
Fraunhofer IPK
D-10587 Berlin, Germany
email; frank-walter.jaekel@ipk.fraunhofer.de

Robert I. M. Young
School of Mechanical, Electrical and Manufacturing Engineering
Loughborough University
Loughborough, UK
email; r.i.young@lboro.ac.uk

Martin Zelm
Standardization Coordinator
INTEROP-Virtual Laboratory
B-1000 Brussels, Belgium
email: martin.zelm@t-online.de

***Abstract* Whilst many ontologies have been developed in research projects over recent decades, we are still a long way short of providing an effective level of interoperability between the systems used, both within and between organisations. It is generally accepted that ontologies and the use of standards are fundamental to achieving interoperability, but effective solutions are proving hard to find. This paper analyses the state of the art in ontological research and standardisation for industry. It highlights the demands of industry and considers the approaches to meeting their needs, providing a perspective on the issues that need to be addressed to provide a step change in interoperability solutions**

Keywords: master data definition; Integration; interoperability; enterprise modelling and ontologies; reference solutions ; international standards.

I. INTRODUCTION

Interoperability is a goal for industry that should provide very substantial benefits in cost saving and improved response times leading to significant commercial value. Without interoperability, communication becomes misleading and the benefits of digitization are lost due to translation errors. This leads to cost and time losses, since, for example, without smooth and interoperable communication between supply chain systems, enterprise resource planning systems and production processing, many manual interventions and lots of interfaces are necessary. This has been known for many years and it has been recognised that both ontologies and standards should play a significant part in achieving interoperability.

While ontologies have been recognised as playing an important part in achieving interoperability, their value to date has been limited. Statements appear such as “it is too complex” or “it is too restrictive” or “how to select the right ontologies or standards for a given question or task”. The problem with using ontologies to date has been that current ontologies have mainly been developed independently. This means that incompatibilities are effectively built-in as definitions of terms are produced independently with no real thought to ensuring compatibility. To do this requires a major effort towards providing an ontological infrastructure which can then be exploited by domain specific ontology builders. Tool chains are also required to orchestrate existing tools and allow cooperative work between the stakeholders i.e the domain experts, data designer, IT system providers and process designers, to support collaborative work, consistency, and usability of any developed ontology.

Industrial data standards are fundamentally important if interoperability is to be achieved when required. However, many industrial data standards fail, from an interoperability perspective, because each specific standard is produced with text-based definitions. This means that the definitions are easy to misinterpret when developing other related standards, leading to incompatibility problems. The use of formal ontologies should overcome this problem by automatically identifying incompatibilities between the definitions of terms, thereby enabling these issues to be resolved. However, new standardised methods towards finding common agreement on formal definitions of terms are also required as building consensus is extremely difficult. Without such methods, acceptance of common definitions will not be possible and any routes to effective mapping between alternative definitions will be difficult.

To date, ontological standards, methods and tools fall far short of meeting the demands of industry. This paper analyses the state of the art and elaborates the needs for ontology development approaches in Section 1, whereas Sections II and III address new tool requirements and standardization to support industrial interoperability. Section IV provides the conclusion.

II. Evaluation and State of the Art

The state of the art with regard to ontologies and the related work in the field of international standardization are outlined below.

A. Ontologies

There is still much confusion in industry as to what an ontology is and why it is important. This is compounded by the large number of definitions that exist in academic publications. The early definition being "a specification of a conceptualisation" where "to specify a conceptualisation one needs to state the axioms that constrain the possible interpretations for the defined terms" (Gruber [1], Noran [2]). This by itself is not easy for a lay person to understand. One which is easier to understand but offers a more open interpretation that comes from an international standard (ISO 18629-1:2004). This defines three terms: **ontology as "**a lexicon of specialised terminology along with some specification of the meaning of terms in the lexicon"; **Lexicon as "a** set of symbols and terms"; **Axiom as "**well-formed formula in a formal language that provides constraints on the interpretation of symbols in the lexicon of a language". This allows for ontologies be defined in simple textual form, as per many standards, or even as dictionaries. However, if the terms are defined using axioms then you have ontologies as they are typically intended today i.e. formal ontologies defined using logic where the terminology is computer interpretable.

The importance of ontologies is reflected in the large number of research projects that have been undertaken since that initial definition in 1993. However, it is clear from industrial experience that domain ontology projects typically do not meet the complex knowledge sharing requirements of manufacturing businesses, where all the key aspects of the business must interact. Although reference ontology projects have shown promise in resolving this issue, these too fall short of fully meeting industry requirements (Palmer et al. [3]). Current, large-scale projects, such as the Industry Ontology Foundry and OntoCommons, that aim to provide a standardised ontological infrastructure that can be exploited and used by domain ontology developers, offer new possibilities towards achieving more effective interoperable solutions (Karray et al [5]). These approaches still have many pitfalls to avoid or to overcome before they can be proven useful to industry, not least the balance of cost / benefit that needs to be shown through effective implementation methods and tools. However, most importantly, they do offer the possibility of a standardised infrastructure which companies can specialise to suit their needs and on which systems developers can build their tools and services.

B. Standards

There are a great many standardisation bodies and organisations involved in developing standards, at national, regional, and international levels. The necessary support for business interoperability must, in time, come from international efforts given the largely global nature of businesses and business interactions. However, more local ad-hoc solutions may provide some benefit to smaller businesses focused purely on their own internal interoperability problems.

Some international standards organisations are run as collaborations between national standards bodies such as ISO and ISO/IEC with specific technical committees involved in interoperability related standards, while others operate as a collaboration between interested business organisations such as the World Wide Web Consortium (W3C), the Open Applications Group Inc (OAGi) and the Open Platform Communications (OPC) Foundation.

From experience in ISO and joint ISO/IEC Technical Committees, the following major bottlenecks in standardization have been identified (Kareron [6]):

1. Lack of strong theoretical foundation and sometimes poor knowledge of fundamental academic works, which would be very helpful to edit more sustainable and agile standards for a digital environment,
2. Standards are sometimes developed in an abstract way with insufficient concrete testing of their implementation in industry,
3. Some standards are not simple enough for a handover and adoption by industry
4. The foundations, either scientific, methodological, or ethical are not discussed enough and not shared by the community of experts and targeted practitioners, which conducts to longstanding misunderstandings in the process of development of a standard,
5. The process of developing of standards is very long; experts are often volunteers; there are issues regarding the necessary competencies; at the end the standards are too complex and present data quality issues, which jeopardizes their use in industry.

From an ontology language perspective, the Web Ontology Language (OWL) is developed and supported through W3C while the Common Logic approach is captured in ISO/IEC 24707. Top level ontologies are standardised in ISO/IEC 21838, while industrial data related ontology standards that have been developed through ISO TC184 SC4 include: ISO 18629 [7], Process Specification Language; ISO 20534, Formal Semantic Models for the Configuration of Production Networks; and ISO 15926-2 [8], Integration of life-cycle data for process plants including oil and gas production facilities. ISO15926 is an interoperability standard for the process industry and includes the Work In Progress (WIP) database. WIP is available online and includes technical class descriptions of all the main equipment items, pipe, instruments, buildings, activities and anything else used in engineering, constructing, procuring, operating and maintaining process facilities.

ISO 15926-2 is perhaps the only one being exploited by industry. However, this uses a 4-dimensional approach which is well suited to the flow of fluids through time but not well suited to typical industrial manufacturing processes which operate on a 3-dimensional paradigm. To our knowledge there are no other international standards to date that define ontologies to support effective industrial interoperability.

OPC-UA [10] is an industrial standard which provides an approach as to how to create information models through the use of a unified architecture. These information models call companion specifications and describes the data structures and their relationships. Standardisation bodies can develop a common layer across multiple different separately developed specifications. OPC-UA illustrates the effect of advertising and motivation about standards that are directly usable by industry and can solve pain-points for industry in terms of interoperability

III. Industrial Deamands and approaches

The industry's demands and potential approaches for a common reference for data management and flexible information exchange between the companies' IT systems are outlined in the following sections.

A. Demands

Digitalisation within an organization requires horizontal and vertical interoperability to manage enterprise applications, IT services and related business processes[4]. Initially it may sound simple to have a common master data structure across an organization. In reality however, each enterprise application has its own master data approach and each system, such as a Programmable Logic Controller (PLC), requires specific realisations of at least IDs for real world objects and data. The challenge is to develop a data management approach which provides flexibility in exchanging IT systems as well as supporting a fast and effective interoperability between the IT systems, including the business processes. This leads to requirements such as

- Common reference for data management which could be provided by a company-wide formal ontology,
- Flexible exchange of IT system by conformity to the common reference instead of a high number of unmanageable interfaces between systems,
- Common understanding of the data acquisition and usage across the company,
- Willingness to exchange information with other organisations through an internal common understanding of terms and data structures.

From industry experience on master data management, the challenge is then to find the appropriate way to develop the overall master data management in terms of an ontology and then how to use this. The problem in industry might be just to have common tools to work on a company ontology between consultants, IT vendors, domain experts and research; but which tools should be used? In a current project the consultant creating the ontology framework just used PowerPoint to identify terms. The company proposed Enterprise Architect [11] or, for distributed work, Conceptboard [12]. In addition, Arrow.app [13] was considered. Protégé was found not to be sufficiently easy to use. A common standardized way to support such cooperative working requirements has been identified by modular approaches to develop the ontology. But the modular approach also needs a tool chain for the initial development and the management of changes.

It would be good to have some standard proposal for a tool chain which is well understood. This does not target the tools themselves, but types of tools for ontology developers, ontology data management, etc. In this industry research project, we used arrows.app for modelling and neo4j [13] for data management, before we started with Protégé. Currently we are not sure if this way will be sufficient for the further use of the ontology.

While research is expected to explore alternative solutions, this is not suitable for industry which requires more stringent and standardized approaches. This problem is under consideration within ISO/TC 184/SC 4/TF 2 "SC 4 reference model for industrial data" [9].

B. Approaches

In general, a high number of ontologies exists as well as some frameworks aimed at organising ontologies. This leaves the question of how to navigate these and what is the benefits of using these ontologies in industry. This relates to the problem of providing easy to apply ontology building procedures and identifying the related toolchain. To provide short term solutions for industry, we do not focus on a general overall ontology due to the clear long-term issues involved in defining a generic solution. Therefore, to support industry, we anticipate providing mechanisms to develop for independent ontologies within the factory scope as well as an opportunity to federate them across factories, if needed. This would hopefully lead to a “best practices” approach to do factory ontology development along with the use of “reference ontologies” to interact between organizations through an ontological infrastructure. Approaches such as the Industrial Ontology Foundry [Karray et al 2020] start to establish such a structure across different ontologies.

Alongside ongoing research for solutions, the demand to implement solutions now increases from industry to provide a common information base capturing all their data assets. In fact, this is a prerequisite to providing effective digitalization. An example is the definition of idents (IDs) which represents the relation between a real-world object and its representation in the digital world. It can be used as a handle to get the data from the object and to get the data to the object. Such IDs are used in different enterprise applications SCM, ERP, MES, PLC programs etc. The challenge is that each application defines the ID differently while in the physical world a fixed identification has to be linked to a part, but which ID should be used? We can use serial numbers, but this would require mappings from the serial number into each of the enterprise applications and especially into each new system.

This example illustrates the issue on just one data element but in real systems hundreds of data elements are exchanged. This calls for a kind of reference structure and in a first step at least the enterprise master data might be considered. An ontology approach for such master data can provide clear semantics of terms as well as clear data structures and interrelations across enterprise applications.

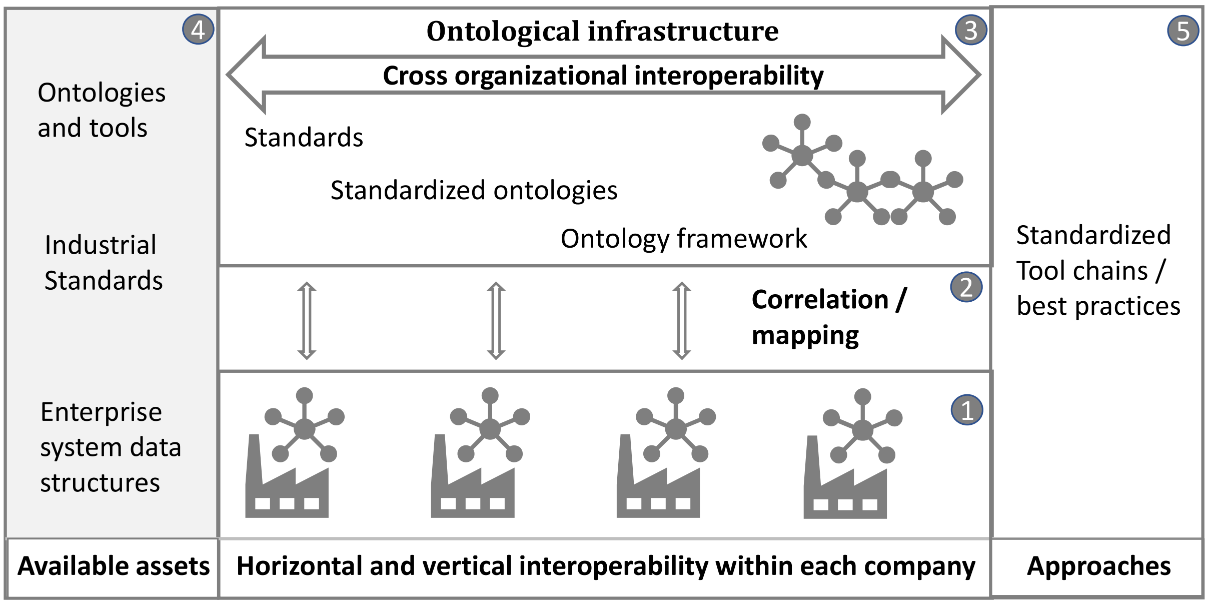

Figure 1: Vision idea influenced by semantic web and experiences from industry

-

Also, the maintenance of the data can be in one place. This would provide a reference data structure for the organization.

Project experiences illustrate that, currently, large companies start to work in this direction.

These industrial demands can be fulfilled independently by each company, but a common approach and consideration of existing standards would be desirable. Therefore, an easy route to identify suitable standards and their use would be helpful. A further challenge is how to support the interoperability between companies as well as with other organizations like government and finance to simplify supply and value chains as well as business networks in general.

C. Vision

The increase demand for digitalisation and a lack of standardised IT solutions that are interoperable by default requires an action to create a common structure of terminologies and relationships. It should ideally be done in a federated way because the necessary definitions for each business depends on company specifics like region, culture, business targets etc. So, in principle each enterprise might create their own common ontology and data structure but would need support in terms of specific approaches and tools to be effective with their definitions. This would give freedom and flexibility to a company in terms of the time horizon because it is not bounded by longer term standardisation processes.

However, in a second step a mapping to a reference framework such as a harmonised set of standards and ontologies could provide a cross organisational interoperability capability [4]. To meet factory scope specific requirements this may not require a generic overall ontology but might be possible with existing industrial standards and sets of ontologies. The federated approach would also allow an independent evolution of the ontologies as well as a different level of maturity in each company. A related vision is illustrated in Figure *1*. It incorporates the following parts:

1. Independent company ontologies to achieve improved internal interoperability e.g. enable company-wide consistent digitalization; provide common master data across the company's IT systems; simplify replacements of IT components. This structure is likely to be private and not public.
2. Separate mappings between the company specific ontology and the required area of the reference structure related to specific scopes. This approach would deliver independence from the standardization time horizon for the companies.
3. Ontological infrastructure and a reference structure for agreed interoperability across companies/enterprises.
4. Available standards, ontologies, information design tools and enterprise data
5. Approaches for the achievement of a company ontology with related methods, tools and best practises as well as mechanism to extend the ontological infrastructure

IV Discussion and Conclusions

The building of a formal ontology as a pre-requirement for standardization of ontologies derived from the collaboration projects seems to include the following issues: 1. Definition of Ontology domains 2. Multi-ontology compatibility 3. Tools and methods to support the development of reference ontologies and exploitation.

The paper identifies the needs and benefits of ontologies from the industry users' point of view, identifying short term focused needs of individual companies versus the more generic longer-term approach of standardisation initiatives. Further, the paper presents a high-level vision of the interaction between ontologies and standardisation.

At present, there exists a number of individual ontologies within a given data space or scope such as a factory or company. As a short-term interim solution to cope with the above issues, a best practice approach would be appropriate through the federation of existing ontologies. Once an ontology exists, the formal definitions of terms within it also exist. When a conflict of compatibility is identified, a cyclic revision process to align the given ontologies will be required. This raises further issues related to the ease or even the possibility of alignment. However, over time, such progressive alignment should be more easily achieved with the development of reference ontology infrastructures which, in turn, should pave the way for new standards to enhance data exchange and interoperability.

Further work of collecting, structuring and harmonization ontology research work results as well as new research is required to elaborate an overall reference infrastructure.

References

[1] T. R. Gruber: A translation approach to portable ontology specification. *Knowledge Acquisition.* 5(2), pp. 199-220 (1993)

[2] O. Noran: Towards a Sustainable Interoperability of Standards. Information Control Problems in Manufacturing, 14(1), pp. 1272-1277 (2012)

[3] C. Palmer; Z. Usman; O. Canciglieri Junior and A. Malucelli and R. M. Young: Interoperable manufacturing knowledge systems, International Journal of Production Research, DOI: 10.1080/00207543.2017.1391416 (2017)

[4] M. Born; U. Greiner; S. Lippe; R. Ruggaber; T. Kahl and J. Ziemann and F.-W. Jaekel: ATHENA Framework for Cross-Organizational Business Processes from an EU Perspective, MIT Press, (2009)

[5] M. H. Karray; N. Otte; D. Kiritsis; A. Farhad; K. Boonserm; C. Willf and R. Aristag and R. Raih and B. Smith: The Industrial Ontologies Foundry (IOF) perspectives, Proceedings of the I-ESA 2020 Workshops, CEUR-W S Vol 2900 (2020)

[6] Y. Keraron and K. Hribernik: Interoperability for Maintenance, Proceedings of the I-ESA 2020 Workshops, CEUR-WS Vol 2900, (2020)

[7] ISO 18629-1:2004; Industrial automation system and integration -- Process specification language: Part 1: Overview and basic principles

[8] ISO 15926-2:2003: Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities

[9] ISO/TC 184/SC 4/TF 2: Technical Report, titled "SC 4 reference model for industrial data", ISO (2022)

[10] OPC-UA, https://opcfoundation.org/, viewed on August 30th 2021. Accessed June 20, 2022

[11] Enterprise Architect, https://www.sparxsystems.com/. Accessed June 20, 2022

[12] Conceptboard, https://conceptboard.com/de/. Accessed June 20, 2022

[13] Arrows and neo4j, https://neo4j.com/labs/arrows/. Accessed June 20, 2022