Towards a method for harmonizing information standards

Xenia Fiorentini, Sudarsan Rachuri, Steve Ray, Ram D. Sriram

National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

Abstract—Designers and engineers use various engineering authoring tools, such as CAD, CAE, and PDM, to generate information objects (engineering objects). On the business side, enterprise level business process modelers use various business authoring tools, such as ERP, CRM, and LCA, to generate information objects (business objects). These information objects, both engineering and business, are represented using information standards. These standards are used to exchange information about engineering and business systems for enterprise level interoperability. One of the main problems of designers, engineers and process modelers is the selection of appropriate standards for interoperability. To ensure enterprise level interoperability, it is absolutely critical that information standards are compared and harmonized as there are overlapping and dissimilar standards available. In this paper, we will sketch a method towards comparing and harmonizing standards based on: 1) informal approach, 2) typology of standards, 3) use-case scenarios, and 4) ontologies. The method is explained using some engineering and business standards.

I. INTRODUCTION

Currently companies need to manage their product information in an integrated, collaborative, and cooperative manner to guarantee an efficient control of their product lifecycles. The product information includes both meta-data about the product, such as design ownership and manufacturing time, and data of the product itself, such as product color and shape. Both kinds of information need to be generated, exchanged, stored, and retrieved consistently by all the stakeholders involved throughout the lifecycle of the product.

Information technologies play an important role in producing information objects, both engineering objects and business objects and enabling interoperability across the networked extended enterprises. In producing these objects, the producer uses different applications to process (authoring and exchanging) the product information: CAD (Computer-Aided Design), PDM (Product Data Management), CRM (Customer Relationships Management), ERP (Enterprise Resource Planning) and LCA (Life Cycle Assessment) systems represent just few examples of these applications. The product information needs to remain consistent, complete, unambiguous and accessible while being processed by these different applications. Interoperability of these application and harmonization of their product information representations are then a prerequisite for the integrated management of the product lifecycle.

The role of product information standards is critical for interoperability. In this paper, we are only focusing on standards that have a formal model that represents product information. The models contain an abstraction of the product information, usually organized in taxonomies and networks of concepts. The applications then instantiate these models by providing the specific information regarding the product being designed, manufactured and/or disposed. The product information, encoded in these standard models, is then easily shared by two or more different applications.

The number of information standards currently available is large and their developing organizations vary from international to local, from large to small, from accredited to informal. When many different standards are available, enterprises find it difficult to choose the appropriate standard and apply it. As an example, the PDM Schema [1], STEP AP214 [2], STEP AP239 [3], OAGIS [4], and OMG PLM Services [5] represent just few of the standards available for the exchange of PDM information between heterogeneous systems. All these standards, although overlapping, are different in content, language and processable expressiveness [6]. First, at the content level, the information encoded in each standard can belong to different domains. As an example, the STEP AP 203 is used within the design domain while OAGIS is within the procurement domain. Second, at the language level, the symbols, conventions and rules used to encode and express the content can vary. As an example, the information models of STEP are developed in EXPRESS [7] while OAGIS in XML [8]. Third, at the expressiveness level, the languages mechanisms that support machine understanding and semantic interpretation can be different. As an example, the XML models in OAGIS are pure syntactical standards, while the EXPRESS models in STEP include some semantics.

Since there is a multitude of information standards, a method to compare and harmonize them is needed. A comparison method could help to understand what the gaps and overlaps between any two standards. The two standards could be harmonized in order to enable interoperability between the different applications that implements these standards. In this paper, we use the word “harmonization” to mean a formal mapping between the concepts and relationships represented in each information standard.
This harmonization is of a particular interest especially when the applications are used to generate different kind of information: engineering related and business related. Engineering authoring tools, such as CAD, CAE and PDM systems, generate what we call engineering objects, while business authoring tools, such as ERP, CRM and LCA, generate what we call business objects. The examples we choose to demonstrate our approach regard the harmonization of a standard describing engineering objects, i.e., OMG PLM Services [5], and a standard describing business objects, i.e., the Open Application Group Integration Specification (OAGIS) [4]. Both standards are widely adopted for representing product information and several enterprise integration problems could benefit from this harmonization.

In section 2, we briefly describe which mapping approaches have been applied in the literature to harmonize information standards. For each approach, we provide a reference to an application example. In section 3, we then sketch our approach towards comparing and harmonizing information standards. This approach includes the following four steps, first the standards are compared using an informal approach, followed by this, the standards are categorized using typology of standards, as the third step the standards are compared based on the application scenarios, and finally they are harmonized by formally mapping the underlying ontologies of the standards. Our conclusions are reported in section 4.

II. BACKGROUND

Efforts towards harmonizing information models have been reported in several application areas. For example, in the biomedical area, the Biomedical Research Integrated Domain Group (BRIDG) project aims to harmonize existing standard information models for clinical research from a variety of sources [9]. To achieve this goal, the group created a semantically conceptual model that was intended to harmonize and unify existing information standards. In the legislation area, as another example, the European Program for an Ontology based Work Environment for Regulations and legislation (E-POWER) project [10] explored the possibility of harmonizing the legislations developed from different European tax authorities. Within this project, a method to formalize the legislative laws in ontologies and to compare them was suggested [11]. In the product information exchange area, many projects within the ISO have attempted to guarantee the interoperability between different ISO 10303 Application Protocols (APs) in overlapping areas (e.g. between AP203 [12] and AP214 [13] for 3D mechanical design). A manual approach is mostly used to discover gaps and overlaps among these APs [14].

III. OUR METHOD TOWARDS INFORMATION STANDARDS HARMONIZATION

Product information standards are usually composed of an introduction, some use-cases (or scenarios), a thesaurus, and an information model. The introduction summarizes the scope of the standard, together with its architecture. Some use-cases then follow the introduction and explain the application of the standard to specified domain. The role of the use-cases is to explain how the standards can be used. The thesaurus provides the definitions of the concepts used in the standard in natural language: the level of formality of these definitions and their descriptions vary from standard to standard. The reader then interprets these definitions depending on his/her background, and uses this interpretation to understand the information model. The model describes the relationships between the concepts and is usually represented in a modeling language, such as UML [15], XML [8] or EXPRESS [7]. The formal semantics of the model, and therefore its interpretation, is dependent on the expressivity of the language used to represent it.

Typically, any standard documentation is composed of hundreds, if not thousands, of pages, which contain introduction, use-cases, thesaurus, and information model. Reading and understanding the documentation is therefore a time-consuming process. This process becomes particularly long especially when there are two standards which need to be not only read and understood, but also compared and harmonized.

To optimize the time and resources allocated to the process of harmonization, and to improve its formality, we propose a method. The steps of this method are as follows: 1) comparison based on informal approach, 2) comparison using typology of standards, 3) comparison based on use-case scenarios, and 4) comparison using formal ontologies. We dedicate this section to the description of each step: we describe our approach and how it is being implemented in our experience with the harmonization of the OMG PLM Services and OAGIS standards.

A. Informal approach

The harmonization process begins with a comparison of the standards: depending on the amount of overlap found, the two standards can be harmonized. The approach adopted in this first step is defined as informal because it does not prescribe any formal method for its realization. The outcome of this step is a general understanding of the standards and a preliminary evaluation of the overlaps between them.

The way this first step is performed strictly depends on the knowledge and background of the person performing the harmonization. As an example, we report here the characteristics of the standards that have been of main interest for the harmonization of OMG PLM Services and OAGIS.

---

1 For the purpose of this paper, the terms scenario and use-case are interchangeable.
The first of these characteristics is the scope of the standards, usually summarized in the first few pages of their documentation. Usually, an overlap in the scopes would normally lead to an overlap of the information models. The first few pages of the documentation also report the relationships between the selected standard to other standards. A list of terms most commonly used in each standard and a research on some implementation cases can help understanding the standards domain. The organizational structure of the standards and the principles used to design it are also other important characteristics: such a structure similarity is a proxy for the evaluation of the harmonization process complexity.

Harmonization: Informal approach

The scope of OMG PLM Services is to represent engineering objects and to integrate them with service technologies. To achieve this scope, the OMG PLM Services specification provides i) a model to structure the engineering information of a product and ii) a model to query and navigate the structure. These two models are called, respectively, the informational model and the computational model. The specification defines a Platform Specific Model (PSM) applicable to the web services implementation defined by a WSDL specification, with a SOAP (Simple Object Access Protocol) binding, and an XML schema specification. The services derived from the combination of the informational model and computational model allow creating, reading, updating and deleting engineering objects.

The OAGIS standard aims to achieve interoperability between disparate enterprise-wide business systems by standardizing the architecture of the messages they exchange. These template-based messages, called Business Object Documents (BODs), are defined in XML and exchanged between software applications and/or databases. The OAGIS Release 9.2 specification includes the description of over 200 BODs.

Results: Informal approach

As a starting point for the harmonization of the two standards, the team decided to focus on the mapping between the nouns of the OAGIS BODs and the informational model of OMG PLM Services. Unfortunately, even though both the standards are based on XML, their XML schemas and XML schema files (.xsd) are differently organized. From the informal analysis, it was clear that the two standards are different not only in their organization, but also in their scope and point of view. STEP with its spin-offs (OMG PLM Services included) and OAGIS model different aspects of product data. STEP models mainly focus on Engineering Objects (EOs), which are outputs of design authoring tools (e.g., CAD and PDM) and generally include information such as product geometry and tolerances.

OAGIS nouns mainly focus on Business Objects (BOs), which are outputs of business authoring tools (e.g., ERP and SCM) and generally include information such as product location, cost and sales order.

Since EOs and BOs usually complement each other, the challenge to the team was to find the product data that travel across engineering authoring tools and business authoring tools. The team found that the intersection between the engineering and the business domains is constituted by the following OAGIS nouns: EngineeringChangeOrder, BOM, and ItemMaster.

B Typology of standards

The second step is more formal than the first one: it prescribes to identify the typology of standards. Categorizing the standards according to their typology helps understand their overlap. As described and elaborated in [16], several typologies of standards relevant to PLM support have been proposed:

Here we summarize, in the form of table (Table 1), the results of this step as applied to our case.

Harmonization: typology of standards

OMG PLM Services was conceived to be used in the engineering domain, where product models are planned and designed. This standard was developed by OMG, which is a computer industry consortium.

OAGIS was developed by a private organization that develops interoperability standards for enterprise-wide business functions, such as ERP, e-commerce, CRM (Customer Relationship Management) and financial. In Table 1, we categorize the two standards based on the typologies previously defined.

<table>
<thead>
<tr>
<th>Category</th>
<th>OAGIS</th>
<th>OMG PLM Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>Content standard: product information modeling and exchange</td>
<td>Content standard: product information modeling and exchange</td>
</tr>
<tr>
<td>Stages</td>
<td>production, use, and identification</td>
<td>Design</td>
</tr>
<tr>
<td>Origin</td>
<td>Open standard</td>
<td>Open standard</td>
</tr>
<tr>
<td>Development process</td>
<td>Consensus standard (private organization developed)</td>
<td>Consensus standard (consortium developed)</td>
</tr>
<tr>
<td>Intent</td>
<td>Interoperability standard</td>
<td>Interoperability standard</td>
</tr>
</tbody>
</table>

Results: typology of standards

A comparison of the columns in Table 1 shows that the two standards mainly differ for the stages of the product lifecycle that they cover. OMG PLM Services was conceived to be used in the engineering domain, where product models are planned and designed. As a consequence, the engineering objects described in OMG PLM Services mainly represent product models. OAGIS has a different prospective, covering other domains such as manufacturing, logistics and sales, and the information of real world products is the main focus.
Therefore, the approach of OMG PLM Services is top-down: the informational model first defines a product model and then links it to actual products (or real world products.) On the other hand, the approach of OAGIS is bottom-up: the information contained in the selected nouns mainly represents actual products and contains references to product models.

The intersection between the information represented in the two standards consists of product engineering aspects that belong to real world products and product models. In OAGIS, the metadata and data of both the real world product and its model are flattened in a unique noun. In OMG PLM Services, the informational model mainly represents both the metadata and data of product models. Therefore, understanding the relationship between data and metadata of product models and real world products becomes the focus of the team.

C Scenario-based approach

Scenarios in information standards have two different roles. First, they help the developing standards committee to reach agreement on the scope and boundaries of the standard. Second, once the standard has been developed, they help the standards users to understand those scope and boundaries. Both these roles are exploited in the scenario-based approach.

The outcome of this approach is composed of an analysis of the scenarios provided in each standard and of a realization of a scenario (one or more) describing when and where the harmonized standard will be used. The analysis of the provided scenarios helps the harmonization team to acquire a deeper understanding of the standards to be harmonized, while the realization of a harmonization scenario helps the team to understand the integration issues.

In the simplest cases, scenarios are composed of an overview, which provides a general description of the scenario, and of a diagram. In the diagram several actors, both humans and software, exchange information necessary to realize the scenario. Well developed scenarios can contain, for example, the description of the actors’ roles, information workflow, assumptions, constraints, exception handling, etc.

Many factors affect the analysis and comparison of scenarios, e.g., their purpose, their actors, the granularity of the information exchanged and the completeness of their description. Naturally, the scopes of the scenarios to be compared should be related, i.e., either they should be similar (e.g., a scenario on purchase order exchange can be similar to a scenario on …), or one of them should be contained in another one (e.g., a scenario on ERP information exchange could include a scenario on purchase order exchange.) Once two related scenarios are found, the similarity between their actors should be analyzed. Scenarios with similar scope and actors should contain similar information. Although the information is similar in the two scenarios it may be defined at different granularity levels, i.e., it has different levels of aggregation. In this case, to make the scenarios comparable, the team could choose either to detail the information exchanged in the scenario with higher level of aggregation, or to aggregate the information exchanged in the scenario with finer granularity. This choice could depend on the background of the team members and on the completeness of the scenario description.

Harmonization: scenario-based approach

The OMG PLM Services 2.0 specification describes 27 use-cases (scenarios) to illustrate the context and way of usage of the informational model: these use-cases fulfill the requirements analysis reported in the PDTNet project [17]. Since all the use cases belong to the engineering domain, their scope mainly regards the exchange and browsing of assembly data, products structure, configuration data, and engineering change request. Naturally, the actors of all the use-cases are the people and software interacting in a PLM System, i.e., a PLM Client, a PLM Server, and a user.

The OAGIS Release 9.2 specification includes the explanation of 61 integration scenarios, in which two or more integrated applications exchange BODs messages. These scenarios can be used as example or modified depending on the needs of each organization. Inventory, ERP and order management are examples of the typology of systems involved in the scenarios. Most of the scenarios belong to the business domain, i.e., they cover business areas such as order management, product data collection, manufacturing and ERP, invoice matching and supply chain integration. The only scenario described in the engineering domain is Scenario 49: Engineering Changes.

Results: scenario-based approach

A comparison between the scenarios of the two standards is hard to achieve since they strongly differ in scope and granularity. While the OMG PLM Services use cases model the engineering aspects of product related data, the OAGIS scenarios model the business aspects. Moreover, the data needed to realize the OMG PLM Services use cases have a fine granularity, while the data needed for the OAGIS scenarios have a high level of aggregation.

The most suitable scenario for this challenge is constituted by the Engineering Change Management (ECM.) The communication, proposal and acceptance of such changes represent the bridge between the manufacturing world and the engineering one. In this scenario, during the manufacturing stage of products, some alterations to the product model need to be made. These alterations will affect the representation of the product model in both the engineering authoring tools (using OMG PLM Services) and the business authoring tools (using OAGIS.)
The OAGIS Scenario 49 seems to address the requirements for the integration scenario of this project as it is focused on engineering information and it describes the integration of engineering and manufacturing systems, i.e., PDM and ERP systems, respectively. For the sake of the project, the team restricted this scenario to the exchange of the BOM messages between design engineering systems and manufacturing engineering systems.

In this restricted scenario, a PDM system conforming to the OMG PLM Services standard exchanges the BOM information with an ERP system. While the PDM system contains the product EBOM, the ERP system needs the product MBOM to generate the manufacturing planning and the purchase orders. The understanding of the relationships between EBOM and MBOM becomes then part of the fourth step.

D Ontology-based approach

In the existing information standards, the terms and the concepts used are mostly in natural language form, even though there is a formal model of the information defined using a specific language. For example, in STEP, the notion of “product” is only understood by humans, but cannot be processed semantically to mean the actual notion of a product. The language chosen to represent that model also restricts the information model. For example, EXPRESS, used in STEP, has limited semantic capabilities.

Doing an ontological analysis of existing standards to identify ambiguities, terminological differences and semantic mismatches will enable to ontologically compare different standards. Examples of such analysis include the OntoClean methodology [18]. Defining richer and more fully axiomatized ontologies would enable capturing formally the knowledge that is embedded in the standard and mapping different standards. In the ontologized version of standards, rules and constraints in standards are formally defined (i.e. using a computer interpretable form of logic), to reduce the ambiguity that is possible with natural language alone [19].

We also need to explore the use of existing ontologies to support the integration of existing standards. This may lead to the identification of new ontologies that serve as mediators between the ontologies associated with each standard. We can pose this problem mathematically as follows. In the following discussion, we use model to denote the information model of the standard.

Let \( M = \{ e, op, r, c=F(e) \} \) denote a model, where \( e \) represents the elements (concepts) defined in the model, \( op \) represents the operations defined acting on the elements \( e, r \), represents relationships between the elements \( e, c \) represents the constraints to be satisfied by the elements \( e \). The elements \( e \) are defined to be belonging to a set \( O \), where \( O \) denotes the ontology for the specific domain \( D \), which \( M \) is modeling. The \( O \) is formal explicit specification of shared concepts and it could be based on some base ontology \( B \). We denote \( M(O) \) to mean that the model \( M \) is based on the ontology \( O \). The problem definition is not complete in any sense. More research is needed in understanding this problem. We are merely sketching the problem of standards interoperability using formal ontologies.

Let \( M(O1) \) and \( N(O2) \) are models of a set of concepts, where \( O1 \) and \( O2 \) represent two different ontologies both based on the base ontology \( B \). The harmonization problem could be posed as defining a mapping \( T: M(O1) \rightarrow N(O2) \) such that

\[
\begin{align*}
1. & \forall e_M \in M(O1), T(e_M) \in N(O2) \quad \text{(Range of } M(O1) = N(O2) \text{). This is required to ensure that there are no missing elements in } N \text{ representing a similar concept defined in } M. \\
2. & \forall e_M \in M(O1), e_M = T(e_M) \in N(O2) \quad \text{(The element (defining a concept) defined in } M(O1) \text{ is equal to the element (the similar concept) defined in } N(O2). \\
3. & \forall e_M,f_M \in M(O1), T(e_M) \circ_N T(f_M) \in N(O2) \quad \text{(Closed under }) \circ_N \text{ )} \\
4. & \forall e_M \in M(O1), e_M \circ_N f_M \in T(e_M) \quad \text{(Conserve the relationships) } \\
5. & \forall e_M \in M(O1), e_M = F(e_M) = c_M = G(T(e_M)) \quad \text{(Conserving the constraints) } \exists \\
6. & \forall e_M \neq f_M \in M(O1), T(e_M) \neq T(f_M) \in N(O2) \quad \text{(Unique elements mapped to unique elements) } \exists \\
7. & \forall e_M \neq f_M \in M(O1), T(e_M) = T(f_M) \in N(O2) \quad \text{(Similar elements mapped to similar elements) } \exists \\
8. & \forall e_M \in M(O1), F(T(e_M)) = T(F(e_M)) \quad \text{(Also } F \text{ is defined for elements in } M, \text{ so will hold for elements in } N) \quad \text{Refer 5) } \exists
\end{align*}
\]

To analyze and solve this problem, many questions need to be answered, and it requires a community of experts to work on these questions. For example, if such a mapping has found where does it belong to, say a domain of mapping called \( S_M \). What are the characteristics? Do define such a domain \( S_M \), we need to define what we wanted to call Standards Interoperability Framework

Harmonization: ontology-based approach

In order to harmonize the selected standards, the team needs to create a BOM reference ontology, to which both the PLM Services schema and BODs schema could be mapped. The BOM reference ontology should represent both the engineering and manufacturing BOM, and should incorporate information about both product models and real world products at different levels of granularity. This requires a clear and formal model of BOM. This ontology could be organized in a multi-layered architecture similar to MOF (Meta Object Facility) [20]. The concepts in each layer and the relationships between layers could be semantically defined. Semantic technologies could enable the harmonization: for example, languages based on description logic could enhance the formalization and allow the consistency checking between schemas and instances [21]. Once the PLM Services and BODs schemas are mapped to the appropriate layer of the BOM reference ontology, the direct mapping between PLM Services and OAGIS can be accomplished.

To realize this, the team would need, as a first step, to analyze the concepts and requirements for the BOM reference ontology, paying particular attention to the representation of product models, actual products and their relationships. As a second step, the team would need
to select a language expressive enough to represent those concepts. OWL-DL [22] or RDF [23] could be two of the candidate languages. The OntoSTEP [24] and Athena [25] projects could represent the starting point for "ontologizing" the PLM Services and OAGIS standards, respectively. As a third step, the "ontologized" versions of PLM Services and OAGIS should be mapped to the BOM reference ontology. For this mapping, the mathematical approach previously sketched could be followed.

IV. CONCLUSIONS

To ensure enterprise level interoperability, it is absolutely critical that information standards are compared and harmonized as there are overlapping and dissimilar standards available. In this paper, we sketched a method towards comparing and harmonizing standards based on: 1) informal approach, 2) typology of standards, 3) use-case scenarios, and 4) ontologies. The method is explained using some engineering and business standards. More research is needed in understanding this problem.

REFERENCES