METAdoc: a framework for development of meta-protocol documents

RICARDO JARDIM-GONCALVES, ADOLFO STEIGER-GARCAO
Departamento de Engenharia Electrotécnica, UNINOVA-CRI
Universidade Nova de Lisboa
UNINOVA - Quinta da Torre, P2825 Monte Caparica
PORTUGAL

Abstract: - There are available many standard models for data exchange covering a wide domain of business and industrial activities. ISO 10303 (STEP) is an International Standard (IS) releasing several of these standards. Each of them is identified as an Application Protocol (AP), strictly covering one scope of activity. Due to the large number of available APs embracing the major requirements for data exchange needed by one organization, nowadays the development of new APs could be made mostly based in a careful selection and reuse of subsets of the existent ones. At this moment there are emerging proposals for methodologies to support the development and implementation of new APs using such approach. However, APs must be technically documented. These documents include a huge amount of data represented in files with several megabytes of text. Although a significant part of the documentation of a new AP can be deduced from those describing the adopted modules, a methodology that could assist in this major task is lacking. Having such methodology established, a framework could be developed supporting the reuse of the existent standard schemata, avoiding time waste and systematic errors, and enabling data mining and data crossing for better selection of the modules. This paper presents METAdoc, a methodology to support the development of meta-AP documents, foreseeing a complete framework for development of standard APs, from documentation to implementation. The presented work results from the research in many international projects, with a strong contribution from the one done with ISO when developing the new ISO10303 Application Protocol: AP236.

Key- Words: - Standards, Methodology, Application Protocol, Meta-Modeling, Reuse, STEP, Documentation

1 Introduction
Many of the application protocols for data exchange in the scope of the major business and manufacturing activities are available as International Standards. Some are developed in the scope of the International Organization for Standardisation (ISO) or the United Nations (UN). Others are developed at regional and national level, as is the case of those developed by CEN and CEN/ISSS Workshops in Europe.

Established as an International Standard in 1994, the ISO10303 STEP has been releasing a set of Application Protocols (APs), i.e., data models valid to be used in the scope of one vertical application, ready to be adopted by the industry and covering the most important activities of the manufacturing process and product life cycle [1].

Nowadays, some of the major industrial companies in the world are using STEP to help in the interoperability problem of its manufacturing systems. Automotive, aircraft, shipbuilding, furniture, building and construction, gas and oil, are some of these industries [2].

Also, PLib (Parts Library, ISO 13584) is an International Standard developed to support the electronic exchange and storage of catalogues of products and parts in neutral format [3]. Adopting the standard data access and representation mechanisms of STEP, PLib includes a methodology for the design, classification and development of catalogues of parts, defining a complete set of resource components for the implementation of data dictionaries and respective catalogues [4][5].

1.1 ISO10303 STEP: STandard for the Exchange of Product model data
STEP - STandard for the Exchange of Product model data, is an ISO (International Organisation for Standardisation) / TC 184 (Technical Committee: Industrial automation systems and integration) / SC4 (Subcommittee: Industrial data) International Standard (IS) officially identified as ISO10303, for the computer-interpretable representation of product information and for the exchange of product data [6][7]. The objective of this IS is to provide a neutral mechanism capable of describing products throughout their life cycle.

STEP publishes a proposal for a methodology for development, implementation and validation of an open architecture for exchange and share of product data, together with a set of public data models identified as Application Protocols (APs).
During the last years the ISO TC184/SC4 community has been working in several definitions of APs for some of main recognized production system areas, as are the automotive, aircraft, electrical/electronics, shipbuilding, oil and gas and building and construction. Nowadays there are 38 APs registered in ISO.

This standard is mainly contributing for worldwide open systems adopting neutral networking communication for product data exchange between heterogeneous systems, both in-house and with third parties. Also, assists on implementation of system-independent architectures, flexible migration policies, contributing to long-term archiving in paperless and life-cycle maintenance support [8][9].

The complete architecture of STEP can be found at STEP On A Page (SOAP) [1].

2 The problem
One of the typical problems found in today’s organizations is when they intend to integrate their applications with a Neutral Format Platform, and the scope of its applications covers more than one of the available Application Protocols [10].

An example of that is taking place in the Cofurn, SMART-fm and funStep AP-DIS international projects [11][12][13], conducting to the development of the ISO10303 (STEP) AP236, which address the integration of applications related with furniture product data representation and furnishing decoration projects [14].

In STEP, other teams are addressing other scope of implementation and they are seeking harmonization between its own Application Protocol and other APs that should be referenced and integrated between them.

A real example of this scenario is related with the Furniture and the Building and Construction domains. STEP provides one Application Protocol, the AP225 that describes the representation for the space of a room [15]. On the other hand STEP is also providing an AP for furniture product data representation, the AP236.

Thinking in one application devoted for decoration of houses, it needs data representation from the room and from the furniture to be placed inside the space for decoration. A similar situation happens with the applications for Shipbuilding where furniture and appliances should be placed inside the ship to furnish it.

Because a standard for data representation usually cannot cover all the range of activities one application needs to handle, the frequently characteristic of horizontality found in many of the enterprise’s applications implies paying strong attention to the necessity of the integration and cooperation of multiple standard Application Protocols [16][17].

Moreover, in most of the cases the development of a new AP includes in its scope parts of already existent ones. This identified overlap must conduct to an automatic reuse of the existent modules. This will save the effort to develop the sub-models and avoids the risk to have the same subject modeled in different ways, and much probably not interoperable [10].

The adoption of a strategy to help developing and implementing APs that support horizontality oriented applications, reusing, partially or in total, the vertically developed Application Protocols, stimulates the intensive use and extensive reuse of existent standards [8][30][31].

However, the available description of these standard models is mostly available as a text document [18]. Although the document is structured in normalized clauses, the major model semantics are described in unstructured text format. Only the schema models described in EXPRESS language [19] are ready to be directly compiled by computers.

As a current practice, when a new AP needs to be developed, and intends to reuse a subset of already existent ones, the standard document is built adopting a “copy & paste” approach, using exclusively the support of an editor of documents, e.g., MS-Word, Acrobat.

This approach becomes most of the times very complex due to the amount of information to deal, and also a source of critical errors, due to the non-systematic approach adopted. Consequently, information cannot be reused easily, as well decision support procedures like data mining and data crossing.

To have the standard AP documents supported by a framework that allows an assisted handling of the document’s information, is a major requirement to stimulate the reuse and extension of existent standard data models in a systematic and accurate way.

There are some proposals for the development and implementation of multi-level APs based in the reuse and extension of existent models [8][9][10]. Should this proposed framework adopted, a global methodology for the development of multi-level APs can be implemented, covering thus the complete cycle: from documentation to implementation.

3 ISO10303 - AP236, a multi-level AP
ISO 10303 AP236 specifies an application protocol for exchange of furniture product definition and interior design project data among manufacturers,
suppliers, retailers, interior designers, and end users [14]. The exchanged data can include product libraries and catalogues and furnishing projects, including graphical information.

One possibility to develop this AP is to create a new complete schema exclusively devoted for this application.

Another approach could be to consider the immediate reuse of already existent APs, and to integrate and harmonize them using a new high-level schema responsible for this integration task. This new schema will act as a meta-AP responsible for the links between the selected modules from the APs it refers. This approach could be extended building a multi-level integration of selected AP subsets.

3.1 AP236 architecture

The AP236 architecture is based on 3 modules selected from existent schemas, i.e., AP214 subset, AP225 subset and Plib#20. The AP236 Linker module is the new module developed to integrate and map the selected ones [20].

Figure 1 shows in EXPRESS-G based notation the AP236 module integrating the connections and the extensions between those modules. The AP236 inherit the architecture of the AP214, to model the furniture and its catalogue of products. A subset of AP225 is extendend to permit the definition of spaces, like room, for decorations projects. The structure of the PLIB#20 permits the representation of expressions and rules, a requirement for the catalogues of furniture due to its parametric nature, e.g., furniture can be customized in size with the price dependent on its length.

4 SC4 guidelines for developing International Standards

The standing document ISO TC184/SC4 N1217 “SC4 Supplementary directives - Rules for the structure and drafting of SC4 standards for industrial data” provides guidelines for developing International Standards, specifying the requirements for the content, layout, and style for the standards developed by ISO TC184/SC4 [21].

It includes specific clauses giving rules and guidelines to document an application protocol. Table 1 depicts the normalized table of contents for a standard AP document, including the following major clauses:

4.1 Documentation of terms and definitions
This clause includes any application-specific terms used in the introduction, scope, and information requirements clauses.

4.2 Application objects
An application object is an atomic element that embodies a unique application concept and contains attributes that specify the data elements that the object comprises. Application object definitions correspond to the application reference model presented in the AP.

4.3 Units of functionality (UoF)
The Units of functionality (UoF) clause lists and defines the units of functionality of the AP. A UoF is a collection of application objects and relationships that conveys one or more concepts within the application context. The definition of an UoF includes the scope of the UoF, a description of the functionality that the UoF supports, and the lists of application objects included in the UoF. Each application object shall appear in at least one UoF.

4.4 Application Assertions
Application assertions specify the relationships between application objects, the cardinality of the relationships, and the rules required for the integrity and validity of the application objects and UoFs.

4.5 Mapping table
The mapping table shows how each UoF and application object of this part of ISO 10303 maps to one or more application interpreted constructs. The mapping table is organized in: Application element, AIM element, Source, i.e., the number of the corresponding part of ISO 10303, Rules, and Reference path [22].
5 METAdoc: Methodology to support the development of meta-AP documents

METAdoc is a methodology resulting from the research and experience of the authors while leading the development of the ISO10303-AP236 in the scope of many international projects, e.g., COFURN, funStep AP-DIS, SMART-fm. Figure 2 depicts this methodology.

The entry point for this method is the standard ISO AP documents [23] that are mostly available in .doc, .pdf and .html formats, organized following the ISO guidelines ISO TC184/SC4 N1217.

After, these documents are translated to XML (Extended Markup Language) [24], using a DTD (Document Type Definition) conformant with the ISO guidelines, in order to have all of them in a pre-normalized representation.

Two repositories support the METAdoc platform. One that will store the complete information interpreted from the standard AP document, and another that will store the application schema models at a meta-model level, normalized using XMI (XML Metadata Interchange) in this case, for an immediate reuse for implementation when creating new APs.

XMI was developed recently by the OMG (Object Management Group) [25][28] with the intention to provide a common mechanism for interchange of models. It has been adopted by most of the popular tools for data modeling, and bindings for main modeling languages have been defined, e.g., for UML [26][29], or for EXPRESS. In this last case, STEP is developing specifications described by Part25 of ISO10303, ”Implementation methods: EXPRESS to OMG XMI binding” [27].

The core of the repository’s Data Base structure is depicted in Figure 3, described using the ISO10303 EXPRESS-G modeling language. Because the adopted normalized language for model representation in this methodology is XMI, before to
be processed this EXPRESS model is translated to XMI.

Also, the data models in the AP described in EXPRESS, are translated using STEP Part25 to XMI, and stored normalized in the repository for easier access and reuse when selected to be included in the implementation of new APs.

METAdoc methodology proposes the APs and their components, i.e., UoF, AOs, AAs, etc., to be catalogued in standard format for a uniform browse and search.

These catalogues are in this case represented using the PLib standard, adopting a neutral classification system that makes reference to the existent APs and components, enabling a normalized representation and access to the AP’s meta information.

6 Conclusion
Addressing the full scope of information representation an organization needs to manage is an enormous task if reuse of previous work is not made possible. However, as the available work was not planned for easy reutilization, a sound methodology is fundamental to guide further developments.

To have an enlarged knowledge of existing standards can be quite helpful to assist in these new developments.

Today there are some methodologies to assist in the development of new APs based on the selection and reuse of subsets and modules of existent protocols. However there is a lack in these methodologies to support the development of its documentation.

This paper proposes the METAdoc methodology to contribute to fulfill this requirement. METAdoc reflects part of the research effort that the authors pursued during the last years in the scope of several international projects. The development of the ISO10303 (STEP) AP236 has been the proof of concept for METAdoc.

7 Acknowledgements
The authors would like to thank all the national and international organizations that supported and funded during the last years a significant number of international projects in the area of integration of applications for complete interoperable systems using standard-based implicit meta-protocols, which resulted in the methodology presented in this paper. Major organizations are the European Commission, CEN/ISSS, Ministry of Industry of Portugal,

Figure 3 – Repository data structure using ISO10303 EXPRESS-G representation
Portuguese Foundation for Science and Technology, Portugal/USA Foundation for Development, IPQ - Portuguese Standardisation Body, ISO TC184/SC4. Also, the authors express recognition for the project partners and our colleagues that work and contribute substantially in the mentioned projects and in the development of ISO10303 (STEP) AP236.

References:

[17] West, Matthew, Integration of industrial data for exchange access and sharing: A proposed architecture