Requirements Analysis for a Semantic Reference Model of Information Processes Individualization

Dmitry Mouromstev
ITMO University
Saint-Petersburg, Russia
mouromtsev@itmo.ru

***Abstract*—In this paper the author summaries some experience on applying ontology engineering in different applications and propose new requirements for developing a semantic reference model of individualization tasks. The individualization as a notion is closely related to process modelling. As a technological background for building application with individualization features the knowledge graphs are considered.**

I. Introduction

As it is well known, reference model is "an abstract framework for understanding significant relationships among the entities of some environment […]. A reference model is based on a small number of unifying concepts […]. A reference model […] does seek to provide a common semantics that can be used unambiguously across and between different implementations." [OASIS SOA]. In other words, such a model is a domain-specific ontology that facilitate communication and development processes. The core of a refence model is a conceptual model consisting of a set of concepts which allow people know, understand, or simulate a subject the represented system and its environment. In this sense the construction of a conceptual model supposes performing of a conceptualization or generalization process [Davies, Islay, et al., 2006].

When we consider a certain application domain for a reference model as a set of processes described by its purpose and the associated outcomes, we should refer to some Process Reference Model (PRM). PRM typically provides a generic functionality and can be used more than once in different models. So a starting point in creating a process model for a specific purpose will be reuse and customization of a PRM according with a supporting methodology instead of creation a model from a scratch [3].

The key benefits of the utilization of reference process models are a timesaving while design of new models, following best-practice processes used in the industry and adaptation of developed models to the individual needs of a user [Becker and Meise, 2011]. Reference models may be constructed both "top-down" and "bottom-up" (Becker and Schütte, 1997). The first approach implements of generally accepted theories and principles. The second approach is based on use and analyses of real-world data, classification and information retrieval techniques. In the work [Rehse at al, 2016] a new approach to inductive reference model development is proposed. This method is based on an execution-semantic similarity measure and capturing the commonalities of the input models in a behavioral profile. As a result, it becomes possible to derive a reference model subsuming the input models' semantics instead of their structure and reach a higher level of flexibility and applicability to inductive reference modelling.

The next important concept is a process individualization. The problem of individualization is not the creation of the only individual trajectory of the process, but the creation of such an individualization model, which together with the domain ontology and the process model will generate the necessary number of options with the possibility of their adaptation under changing process conditions. Some approaches to a process modelling are considered below.

The final and the most important feature of considered reference models is semantic layer that provides capabilities for reasoning and knowledge base development for individualization purposes. The latest trends here are building so called Knowledge graphs, which include graph repositories of semantic metadata (or knowledge formalized using special formal semantic languages like RDF) and ontologies being as a semi-structured conceptual scheme for a process data. The latter quality distinguishes knowledge graphs from databases, making it possible to perform intelligent data analyses by means of linking, reuse, and enrichment tasks in the data management lifecycle.

II. Related work

A. Reference models

A comprehensive analysis of the concept of the reference model and approaches to its construction and its role in information modelling is provided in [Ahlemann, 2009]. Ahlemann in his work considers a number of models starting from one of the first reference information models for project management in the architecture, engineering, and construction (AEC) industry published by Froese, who called it a ''standard model" [Froese, 1992] and later developments like a combined reference information model for process and project controlling published by Schlagheck [Schlagheck, 2000].

B. Process models

The review of the basic approaches for a modelling if information processes are described in [De Nicola at al., 2007], [Merunka, Vojtěch, 2010], [Rassler, Jochen AT AL., 2008]:

• *Descriptive models* like EPC, IDEF, UML and BPMN. These models lack the systematic formalization necessary to implement the inference mechanism.

• *Procedural languages* like BPEL and XPDL. These languages are not intuitive enough for human use, there is no declarative semantics for the reasoning mechanism.

• *Object-oriented programming* like BORM, PML, BOOPM. Mainly such models are used in software development tasks and not intendent to be a modelling language.

• *Formal languages* like PSL, Pi-Calculus, Petri-Nets are based on strict mathematical principles and difficult to understand.

• *Hybrid models* based on neural networks like Transition-Based Neural Model, Narrative Event Evolutionary Graph Construction (NEEG). These models are computational effective but lack explanatory features.

• *Ontological models* like The Business Process Modeling Ontology and BPAL (Business Process Abstract Language.). Provide a good comprehensive level and being combined with hybrid neural models gives a lot of possibilities.

C. Ontologies and Knowledge Graphs

A detailed review of the use of ontological approach is given in [Mouromtsev, 2020]. The role of ontologies is to provide semantic annotations for real-world data from heterogeneous sources or metadata, which can be represented through ontology concepts. The set of these annotations are linked by object and data properties of the ontology. As a result, it is possible to use a knowledge graph as structured and annotated knowledge base via end-point and semantic queries.

III. THE REQUIREMENTS

A. Data requirements

Usually, the data in information systems and external sources are incomplete, noisy and contain implicit knowledge in the form of text or another human-oriented format. In this case, it is effective to use the methods of machine learning and embedding models that allows to improve data quality by the following operations:

1) representation of raw data from a certain source of information in the form of RDF triples <subject, predicate, object>,

2) annotation, linking and enrichment of data using domain ontologies and specific annotations schema,

3) improving the quality of data using machine learning methods and statistics in the tasks of extracting named entities and relationships, identifying topics, etc.

4) the assembly of the obtained triples of the knowledge graph using the reference model.

B. The architecture of the reference model

The structure of the reference model for the individualization of information processes (fig. 1) includes several layers for modeling concepts of varying levels of abstraction:

1) upper-level abstractions for modeling individual's behavior and their information needs;

2) general concepts of the structure of data sources, as well as models of specific processes in the subject area (for example, smart manufacturing, e-learning, digital budget, etc.);

3) specific concepts for data access and integration in terms of the domain ontology.

The upper two levels are practically independent of the specifics of the subject area and specific information systems. At the same time, the lower level can be adapted to specific requirements or even divided into several sublevels if necessary. Such modifications of the lower level of the reference model can be performed quite often, because the process of individualization itself involves the construction of a separate model for each user based on the data that they use or generate.

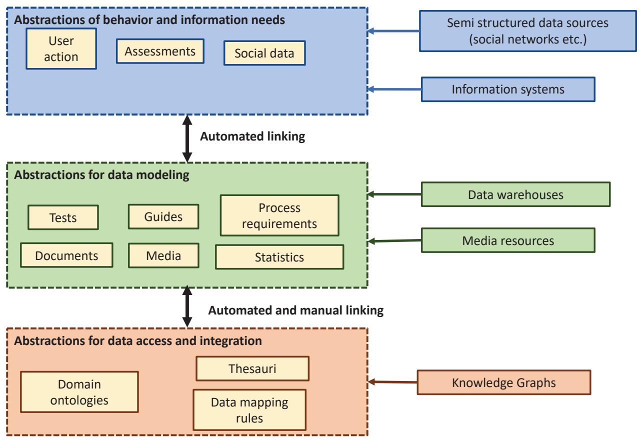

Fig. 1. The structure of the reference model

IV. APPLICATIONS, EXTENSION AND CUSTOMIZATION

A. Smart Industry

The difficulties and barriers of the digitalization of economy and production intellectualization are associated not with a lack of digital services and data, but mainly with a variety of information platforms and data formats. Traditional approaches in ICT do not offer ways to overcome these barriers, but on the contrary, in many cases they make the situation worse when introducing another information system or standards. At practice, the tasks of exchanging data between such systems are solved by the efforts of analysts and data engineers of companies. Obviously, to solve these problems it is necessary to provide automated integration of heterogeneous data sources into a single virtual information environment, which should be based on formalized knowledge. There exist several ontologies for smart manufacturing (ADACOR, MSDL, MASON, CDM-Core etc.). But no one of these ontologies provide features for individualization. Being extended by proposed requirements a new reference model for smart industry may be developed.

A. Document knowledge modeling

Extraction knowledge from a bulk of documents is the one of the most important tasks nowadays. Different standard specifications, guides, laws, agreements and many other types of documents are required for only process in our live. A number of projects related to developing knowledge bases from documents are known. Also, there are many ontologies for knowledge representation in documents like KNORA Ontology for representation the entire document and its usage, Document Component Ontology for modeling the structure of the document, Dublin Core Ontology for a representation of the thesaurus of terms, linguistic ontologies like Lemon Core Ontology or NIF Ontology and many others. But all these ontologies allow represent only static knowledge. For the purposes of individualization, it is necessary to integrate models of documents with user behavior and needs models according requirements described in this paper.

A. Education

Individualization in education is a logical and necessary step in the evolution of e-learning, that should shift education from mass orientation to personalization of learning processes. This change raises many methodological, technological, and conceptual issues, such as modeling of a learner's profile, individual learning paths, individualized assessment of learning outcomes, and some others. There are developed ontologies for education that separately solve the tasks listed above. These are ontologies such as The Academic Institution Internal Structure Ontology (AIISO), The Bibliographic Ontology (BIBO), The Ontology for Media Resources (MA-ONT) TEACH (Teaching Core Vocabulary), FOAF ontology (Friend of a Friend), etc. However, questions regarding implicit or non-declared knowledge in Learning Management Systems cannot be resolved directly using these ontologies. Models of namely such knowledge are necessary for the individualization of e-learning processes and the requirements proposed in this paper allow you to build a reference model for these purposes.

VII. Conclusion

The problem of development reference models is well known task and related issues are considered in the beginning of this paper. Reference models are closely related to concept (ontological) and process modelling. Both are core elements for individualization task. In this work the basic requirements for data lifecycle and architecture of the abstract semantic reference model are presented. Also three application areas are considered for implementation of proposed requirements.

References

[1] "OASIS SOA Reference Model". OASIS SOA Technical Committee. Retrieved 2014-10-13.

[2] Davies, Islay, et al. "How do practitioners use conceptual modeling in practice?." Data & Knowledge Engineering 58.3 (2006): 358-380.

[3] Marne de Vries, Alta van der Merwe, Paula Kotze and Aurona Gerber. (2011) A Method for Identifying Process Reuse Opportunities to Enhance the Operating Model. 2011 IEEE International Conference on Industrial Engineering and Engineering Management

[4] Becker, J. and Meise, V. (2011), "Strategy and Organizational Frame", Process Management. A Guide for the Design of Business Processes, J. Becker, M. Kugeler and M. Rosemann (eds.), Springer, Berlin, pp. 91–132.

[5] Becker, J. and Schütte, R. (1997), "Reference Information Systems for Retail: Definition, Use and Recommendations for Design and company-specific Adaption of Reference Models", Wirtschaftsinformatik, Springer, pp. 427–448.

[6] Rehse, Jana-Rebecca, Peter Fettke, and Peter Loos. "An execution-semantic approach to inductive reference model development." (2016).

[7] Ahlemann, Frederik. "Towards a conceptual reference model for project management information systems." International Journal of Project Management 27.1 (2009): 19-30.

[8] Froese T. Integrated computer-aided project management through standard object-oriented models. Center for Integrated Facility Engineering: Stanford; 1992.

[9] Schlagheck B. Objektorientierte Referenzmodelle für das Prozess- und Projektcontrolling: Grundlagen – Konstruktion – Anwendungsmöglichkeiten. Wiesbaden: Gabler; 2000.

[10] De Nicola, Antonio & Lezoche, Mario & Missikoff, Michele. (2007). An Ontological Approach to Business Process Modeling. 1794-1813.

[11] Merunka, Vojtěch. "Object-oriented process modeling and simulation-BORM experience." Trakia Journal of Sciences 8.3 (2010): 71-87.

[12] Rassler, Jochen, and Reiner Anderl. "PML, an Object Oriented Process Modeling Language." Computer-Aided Innovation (CAI): IFIP 20th World Computer Congress, Proceedings of the Second Topical Session on Computer-Aided Innovation, WG 5.4/TC. Springer, Berlin, 2008.

[13] D.I. Mouromtsev. "Models and methods of e-learning individualization in the context of ontological approach" Ontology of Designing, 10, 1, (2020) 34-49.