

## **THE TECHNOLOGY AND TOOLS FOR THE BUILDING OF INFORMATION EXCHANGE PACKAGE BASED ON SEMANTIC DOMAIN MODEL**

**Yu.M. Akatkin, M.G. Bich, E.D. Yasinovskaya<sup>a</sup>, A.V. Shilin**

*Plekhanov Russian University of Economics, 36 Stremyanny per., Moscow, 117997, Russia*

E-mail: <sup>a</sup> elena@semanticpro.org

This paper presents the technology developed by the authors with the aim to improve the semantic interoperability in heterogeneous environment, where the systems use web-services orchestrated by an object-oriented exchange bus for cross-agency interaction and information sharing. The suggested solution, including the set of tools, allows to map the models of interacting information systems with a unified data model (domain ontology) on the semantic level during the development of an information exchange package.

**Keywords:** Semantic models, Semantic interoperability, Semantic integration, Cross-agency interaction, Information sharing, Domain data model, Information exchange, NIEM, SDMX, SOAP, Eclipse Papyrus.

Yuri Akatkin, Michael Bich, Elena Yasinovskaya, Andrey Shilin

Copyright © 2021 for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 1. Introduction

Heterogeneous environment inherent for information sharing requires the application of integration methods, which will guarantee the achievement of unambiguous and meaningful interpretation of data at all levels. Following to the opinion of expert community the authors consider US National Information Exchange Model (NIEM) [1,2] as one of the most developed and widely used technologies for information sharing during cross-agency interaction. NIEM is a US national standard. This model as well as the methods and technologies for its implementation provide full stack support for developers by using both proprietary software tools (e.g., IBM [1], Oracle<sup>1</sup> и Microsoft<sup>2</sup>) and a set of open-source software tools<sup>3</sup>.

The development of Information Exchange Package Documentation (IEPD) sets the basis for the application of object-oriented approach in NIEM. The developers build IEPDs by implementing all the matching elements from NIEM Core or NIEM domains and where applicable, they can extend this content with new items to meet those information requirements which NIEM does not cover. Thoroughly elaborated over the past 10 years IEPD building process includes 6 steps [1,1]. There is an appropriate toolkit for each step of the IEPD development life cycle.

In object-oriented approach the data models providing semantic interoperability for interacting heterogeneous information systems via are created in compliance with the methods, which include the analysis of exchange business processes in the domain context. However, the built exchange object model obtains only those entities that are essential for this specific information exchange as well as their properties and relationships.

Within the development of the Semantic Web the use of semantic models, actively studied and tested in practice in ontology engineering [3,4,5], has become the dominant method for solving the problem of heterogeneous data semantics.

The authors consider significant to release the potential for the convergence of NIEM-type models, providing data integration on federation principle, with Semantic Web models, supporting semantic annotation. For example, OMG offers such approach in the initiative called the Semantic Information Modeling for Federation [6,7]. The research in this direction had been going on for the last 10 years<sup>4</sup>, although it has not yet received a wide practical implementation. Nevertheless, the enrichment of object-oriented data exchange models with semantics promises the fastest achievement of a new level of semantic interoperability.

## 2. Methods

The authors represent the technology and tools developed following the methods of software design and development adopted in Russia, as well as in accordance with the internationally used Design Science Research Methodology for Information Systems Research (DSRM) [8].

The research of semantic integration methods from the prospective of 8 years' experience, presented in several works<sup>5</sup>, served as the Problem identification and motivation stage, which highlighted the Objectives for the solution, demonstrated in this paper (Section 3.1). The technological stack including the development tools represented in Section 3.2 together with the description of suggested information exchange package (Section 3.3) show the results of Design and development stage. Section 3.3. demonstrates the use case and application of our solution for cross-agency information sharing in chemical and biological security domain. It represents the Implementation stage

---

<sup>1</sup> [https://blogs.oracle.com/xmlorb/entry/oracle\\_niem\\_resources\\_site\\_launches,](https://blogs.oracle.com/xmlorb/entry/oracle_niem_resources_site_launches)

<http://www.oracle.com/us/products/applications/public-sector/niem/index.html>

<sup>2</sup> <http://blogs.msdn.com/b/jrspinella/archive/2011/06/20/biztalk-2010-real-world-example-part-1.aspx>

<sup>3</sup> <https://www.niem.gov/tools-catalog>

<sup>4</sup> <https://github.com/ModelDriven/SIMF/tree/master/Presentations>

<sup>5</sup> <https://www.researchgate.net/lab/Yuri-Akatkin-Lab>

and covers Demonstration и Evaluation activities as well. The authors aim to present the technology and tools to start the discussion during the GRID Conference at Communication stage and use more opportunities for further implementation. Section 4 summarizes current experience and provides demo links for detailed instructions and screen shots for everyone who is interested in testing the suggested solution in real time.

### 3. Solution

#### 3.1 Problem and Objectives

Based on the current experience in the arrangement of cross-agency interaction and conducted research (e.g., [9]) the authors consider that at the present time the topical problem in information sharing is the lack of methods and software tools combining the informative richness of semantic models, describing the interacting information systems domain, with the power of industrial solutions for object-oriented data exchange.

To overcome this challenge, the authors suggest the solution aimed to achieve the following objectives:

- To support mapping of an information system model with the unified data model (domain model) used for the cross-agency interaction;
- To enrich an object information exchange model (exchange model) with semantic descriptions by linking its elements with the domain model, supporting integrity and verifiability;
- To generate object domain models from semantic models automatically.

#### 3.2 Components

The developed solution includes four interrelated components [fig. 1].

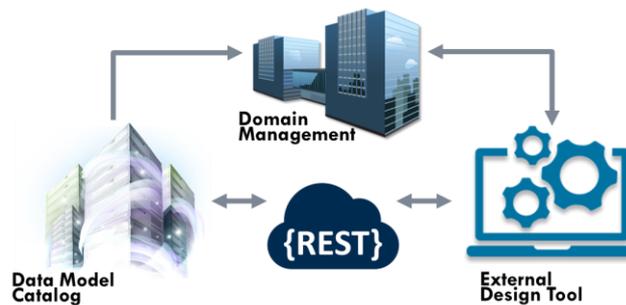


Figure 1. Component Diagram

**Data Model Catalog** provides registration, storage and retrieval of object and semantic models. The authors have implemented Asset Description Metadata Schema (ADMS)<sup>6</sup> extended with the classes and entities to support and track cross relationships of the assets as well as to use domain management services for experts' collaboration [9].

**REST**<sup>7</sup> service provides access to the catalog and its modification. It serves to search an asset search, gets its description (in JSON<sup>8</sup> format) and contents (as files of various formats). Through transactions the service software ensures the integrity of the catalog when creating a new asset or saving a modified version of the existing one.

**Domain Management component** is responsible for joint work of experts on the validation of exchange models as well as on the domain improvement. The component supports the collaboration

<sup>6</sup> <https://www.w3.org/TR/vocab-adms/>

<sup>7</sup> [https://en.wikipedia.org/wiki/Representational\\_state\\_transfer](https://en.wikipedia.org/wiki/Representational_state_transfer)

<sup>8</sup> <https://www.json.org/>

functions. It provides the set of tools for setting up and executing business processes using various teamwork functions (task management, commenting, voting, notification, etc.).

**External Design Tool** serves for the development of exchange models. The authors suggest using Eclipse Papyrus<sup>9</sup> extended with additional plugins: (1) functions for REST service application and domain import; (2) specially developed Papyrus Architectural Context. This architectural context includes: (a) UML profile for exchange model development; (b) icon palette and (c) additional Papyrus plugins. These plugins realize (1) filling special properties of UML classes with IRIs of domain elements, (2) OCL and programmed model validators, (3) automatic conformity control for object and semantic models, (4) the domain ontology file containing OWL and the automatically generated UML domain model. The same IRI in different object models correspond to equal semantics, and automatic control provides equivalent syntax.

### **3.3 Information Exchange Package**

The authors have primary implemented such approach in the Center of Semantic Integration (CSI) project [9]. It is important to highlight the difference between the CSI Information Exchange Package created for cross-agency interaction in our solution and NIEM IEPD.

There are two main features distinguishing the suggested solution:

- The use of semantic technologies during the development and publication of the information exchange package;
- The additional data fixing the collaboration between developers and domain experts (comments, recommendations, discussion).

Considering the general logic of NIEM IEPD, CSI Information Exchange Package includes:

- Ontological representation of data model and proposals for its extension;
- Object representation of data model in the format of the tools used for its design;
- Platform-dependent elements simplifying the implementation of the package into applied information systems;
- Extended package documentation, including semantic annotation for the object models.

### **3.4 Use case**

This use case represents the developed solution in the process of cross-agency interaction in the chemical and biological security domain regulated by the Ministry of Healthcare of the Russian Federation. The fulfillment of the following tasks demonstrates the capability of suggested technology and tools:

- To design information exchange models;
- To build data exchange schemes using SDMX<sup>10</sup>;
- To use the developed schemes for tuning exchange web-services.

Figure 2 represents the contents of CSI Information Exchange Package and the screen shots from CSI web interfaces [Fig.2].

---

<sup>9</sup> <https://www.eclipse.org/papyrus/>

<sup>10</sup> Statistical Data and Metadata Exchange (SDMX), <https://sdmx.org/>

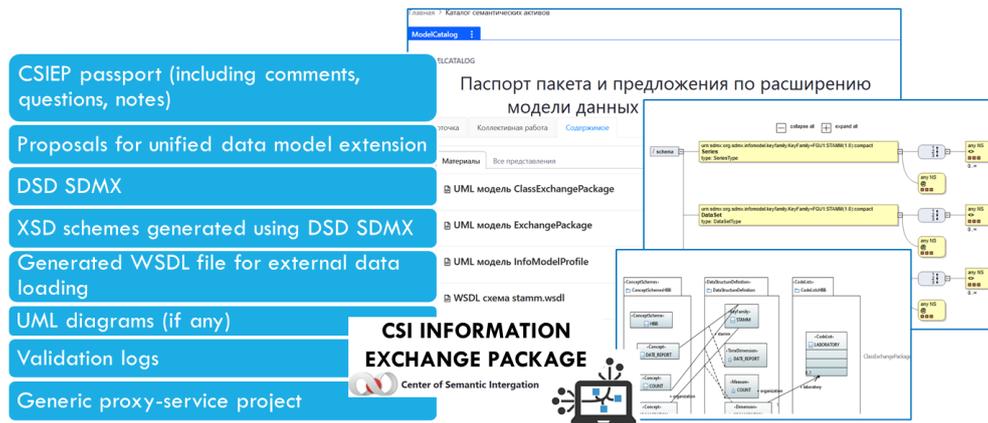


Figure 2. The Contents of CSI Information Exchange Package

To support the lifecycle of CSI Information Exchange Package the authors consider efficient to implement NIEM methodology for IEPD building. The main actors are the developers of agency information system responsible for the information exchange (developers), domain experts, supporting the development of the unified data model (domain ontology), and domain stewards with the authority to modify it.

During steps (1) *Scenario Planning* and (2) *Requirements Analysis*, the developers use their own tools, resources and follow the agency regulations. To work with the suggested solution at the next steps, they get access to web interfaces and download External Design Tool.

At step (3) *Mapping and Modeling*, a developer: (1) imports the current version of domain ontology; (2) creates a new UML exchange model; (3) maps this model with the developed Architecture Context and fills it with typical content, if necessary, or (4) adds his own elements in UML notation. Thus, the elements of the exchange model will have the link to the domain ontology entities associated via indicated IRIs.

At step (4), *Building and Validating*, the developed exchange model gets formal (OCL) and program (embedded program code) validation. If there is no matching element in the domain ontology, the tool automatically creates an extension proposal. It indicates both the suggested element and its place in domain ontology.

At step (5) *Assembling and Documenting*, the tool assembles the CSI Information Exchange Package that archives the results of modeling. As described in section 3.3 it includes information exchange model, extension proposal (if necessary) and supplementary files (UML diagrams, test data, XSD and WSDL schemas, etc.). The REST service transfers the package to the Data Model Catalog.

At step (6) *Publishing and Implementing*, domain experts analyze the received package using Domain Management tools. They check the package and the suggested extension proposals. They can accept or decline them, as well as suggest matching domain elements or making more precise mappings. If the extension proposal contains concepts and data still not reflected in the unified data model, the domain experts can refine or extend it in compliance with the developed information exchange model. After the approval the domain steward, publishes the package in the catalog for further reuse. Notification informs the developers about all decisions, package status and domain ontology updates.

## 4. Conclusion

Semantic interoperability is the point of great importance for the integration of information systems in heterogeneous environment of cross-agency interaction. The authors suggest combining the informative richness of semantic models describing the domain with the power of industrial solutions for object-oriented data exchange. It is necessary to create the technology and software tools to implement this approach and to support the developers of interacting information systems.

This paper demonstrates the solution allowing semantically map the models of interacting information systems with a unified data model (domain ontology) during the development of information exchange package. Indication of IRIs in the object model allows enriching the formal representation of the object structure, for example, in the form of an XSD schema, with additional information. Using the developed models, the special tools can generate exchange interfaces and automatically adjust semantic web services. Presented solution provides an unambiguous correspondence between the object and semantic models throughout the development cycle and ensures the transformation of models, which, in its turn, will serve to implement the semantic service bus for the arrangement of cross-agency information sharing.

All demo links and detailed instructions for the demonstration of presented solution are available on <http://csi.semanticpro.org/material2021.htm>.

## References

- [1] Government Information Sharing & Advanced Insight and Analytics. The National Information Exchange Model (NIEM). IBM Software Information Management, 2010. Available at: <http://www.acceleratedim.com/whitepapers/IBM-NIEM-Whitepaper-Final.pdf> (accessed 03.09.2021)
- [2] NIEM User Guide Vol. 1, 2008. Available at: <https://reference.niem.gov/niem/guidance/user-guide/vol1/user-guide-vol1.pdf> (accessed 03.09.2021)
- [3] Peristeras Vassilios. Semantic Standards: Preventing Waste in the Information Industry. IEEE Intelligent Systems. Vol. 28. P.72-75. doi:10.1109. MIS.2013.115, 2013. Available at: <https://ieeexplore.ieee.org/document/6682939/> (accessed 03.09.2021)
- [4] Walaa, S. Ismail, Mona, M. Nasr, Torkey, I. Sultan Ayman E. Khedr. Semantic Conflicts Reconciliation as a Viable Solution for Semantic Heterogeneity Problems. (IJACSA) International Journal of Advanced Computer Science and Applications. Vol. 4. No.4, 2013. Available at: <https://pdfs.semanticscholar.org/22bb/9c1c27049fce16d1546c54d51e0ede2854d8.pdf> (accessed 03.09.2021)
- [5] Madnick S.E., & Zhu H. Improving data quality through effective use of data semantics, Data and Knowledge Engineering, 59(2). pp. 460-475, 2006. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3747&rep=rep1&type=pdf> (accessed 03.09.2021)
- [6] Semantic Information Modeling for Federation (SIMF) Request for Proposal, 2012. Available at: <https://www.omg.org/cgi-bin/doc?ad/11-12-10> (accessed 03.09.2021)
- [7] Semantic Information Modeling for Federation, EMTECH Semantic Technology & Business Conference, 2011. Available at: [http://ontolog.cim3.net/file/work/OntologySummit2012/2012-03-01\\_Ontology-for-Systems-Federation-n-Integration/OntologSummit2012\\_Semantic-Information-Modeling-for-Federation--CoryCasanave\\_20120301.pdf](http://ontolog.cim3.net/file/work/OntologySummit2012/2012-03-01_Ontology-for-Systems-Federation-n-Integration/OntologSummit2012_Semantic-Information-Modeling-for-Federation--CoryCasanave_20120301.pdf) (accessed 03.09.2021)
- [8] Peffers, K., Tuunanen, T., Rothenberger, M., A. Chatterjee, S. Design Science Research Methodology for Information Systems Research, January 2008, Journal of Management Information Systems 24(3):45-77, 2013. DOI: 10.2753/MIS0742-1222240302
- [9] Akatkin, Yu., Yasinovskaya, E., M. Bich. Semantic Information Management: The Approach to Semantic Assets Development Lifecycle. Proceedings of GRID 2018, Available at: <http://ceur-ws.org/Vol-2267/447-452-paper-85.pdf> (accessed 03.09.2021)