VISUAL INTERFACE FOR ADAPTATION OF DATA SOURCES TO SEMANTIC WEB

Olena Kaykova [1], Oleksiy Khriyenko [2], Mikko Kovalainen [3], Andriy Zharko [4]

Department of Mathematical Information Technology,
University of Jyväskylä, P.O.Box 35 (Agora), 40014, Jyväskylä,
FINLAND

*1helen@it.jyu.fi, *2olkhriye@cc.jyu.fi, *3mipeko@mit.jyu.fi, *4anzharko@cc.jyu.fi

ABSTRACT

Development of computer communications has catalyzed integration processes in world communities. Initial efforts concerning integration have showed great heterogeneity between data formats being exchanged between applications. One of the aspirations of Semantic Web Activity lead by W3C is provision of interoperability on concept level between software systems. The idea of Semantic Adapter presented in this paper continues efforts in this direction. It assumes development of software that would "wrap" data entities of legacy systems in additional semantic markup based on common ontology. Since only human knows the semantics of the data provided by data sources, there must be a handy visual interface for binding data source and metadata notion. The chosen way of Semantic Adapter configuration will be intuitively understandable for users, who don't have knowledge about Ontology presentation standards. Semantic Adapter described in this paper is one of the constituents of GUN - *Global Understanding eNvironment,* being developed by Industrial Ontologies Group [11].

KEYWORDS
Visual interface, Ontology, Semantic Web, Semantic Adapter.

1. Introduction

Very few business applications can live in isolation. Most often, applications have to be integrated with other applications inside and outside the enterprise. This integration is usually achieved through the use of some form of "middleware". Middleware provides the "plumbing" such as data transport, data transformation, routing etc. [1]. Enterprise integration can also help to reduce costs, increase operational efficiencies, expedite time to market, and improve return on information technology investments. Without enterprise integration, various infrastructures will lack the robustness and flexibility in a dynamic economy [2]. Now integration is the direction of success for enterprises. They invest large portions of their ICT budget on informational integration.

The recent trends in industry connected with enterprise integration demand solid technology to provide interoperability between heterogeneous applications interacting within and across industrial enterprises [3]. Applications of Semantic Web in recent researches [3-7] have shown many useful features for this domain. Semantic Web Activity [9] lead by W3C [8] provides support in standards for Semantic Web (RDF-Schema, DAML+OIL, OWL, DAML-S). The scale of Semantic Web Activity widens attracting more and more parties.

Companies developing software can contribute much in integration processes providing interoperability for legacy systems that now have to be integrated. This can be done through development of software plug-ins (adapters) that has on the one hand software-specific interface and on the other hand standardized interface. Original data in order to be available to external applications needs to be presented in some format, which those applications can process.

Making things presented in commonly understandable form is one of the goals of Semantic Web technology. Interoperability issues can be overcome if software-developer companies utilize the potential of this technology. Software systems can be adapted to produce data using common "language" provided by Semantic Web approach.

Semantic Web's ontologies can be used as manifestation of data representation standards and configurable *semantic adapters* can be developed to support various standards of data representation and transform them into common form of semantically annotated data. This will enable data retrieval by applications, which are developed with support of common (semantic) representation, from heterogeneous data sources that are of different formats, standards, data schemas, etc.

The paper is organized as follows. In Section 2 the idea of Semantic Adapter is presented as context for the next, main section. In Section 4 the idea of human visual interface for Semantic Adapter configuration is stated. Finally, in Section 4 we present our conclusions and point out some directions for further research.

2. Idea of Semantic Adapters

Types of data to be semantically annotated and integrated are not limited only to digital documents and database content. Real world object like human, device can also produce data that are in interest to be exchanged between it and other system, for example, to give directions and orders or represent device state and results of self-diagnostics.

Applications need close connection with resource and ability to operate on it (for resource updates, data retrieval, monitoring, diagnostics, etc.) Taking into account great variety of possible resource types, different in their formats, ways of access to them, and many other aspects (from text file on server to device accessible via hardware interface and human accessible via SMS), specific *data source connectors* are used in semantic wrappers. The same semantic wrapper component can be used for different types of data retrieving way of the same class of data source (see Figure 1).

Any application can be set up to maintain virtually any kind of supported data sources using semantic wrapper. Development of data source connectors for each class of resource is required for their integration in some environment, for which semantically annotated content will be produced from these resources.

Ontology-based approach stands as an alternative for development of domain-specific set of standards/vocabularies/procedures for information exchange. We use ontology concept and data representation framework, which was developed within Semantic Web Group activities [9] for provision of the World Wide Web with the same features as we require: interoperability and data integration in heterogeneous environment. Ontology-based information management is going to be more flexible and scalable, and also it has potential to become next-generation standard of information exchange in the Internet.

Type of software, which uses data being described in an ontological way, can vary depending on needs. It can be a data browser, control panel of operator, computing system, database storage, etc. Because of the way data is represented, it will be never processed incorrectly, since software can check itself whether data semantics, as annotated, is the same or compatible, as data processing unit needs.

Additional benefit comes from data annotation for software development even if there is no need to deliver information outside of origin computing system: no more needs to develop special formats of domain-specific data exchange between application, since it is already presented in common standard by means of ontology. Software can be developed in a modular, scalable manner with support of this standard (ontology). Such commitment to the shared (upper-) ontology will provide compatibility of software.

Extensibility of ontology is its inherent feature. As new types of information and needs to present them appear increasingly, ontology can be extended to reflect those changes and used further as a standard.

Development of standardized information for domain-specific systems requires:

Phase 1: Building ontology;

Phase 2: Adaptation of data sources for producing ontologically annotated data;

Phase 3: Development of ontology-based data processing tools.

Ontology engineering phase includes development of upper-ontology (schema for ontology) and development of ontology itself, which includes domain-specific

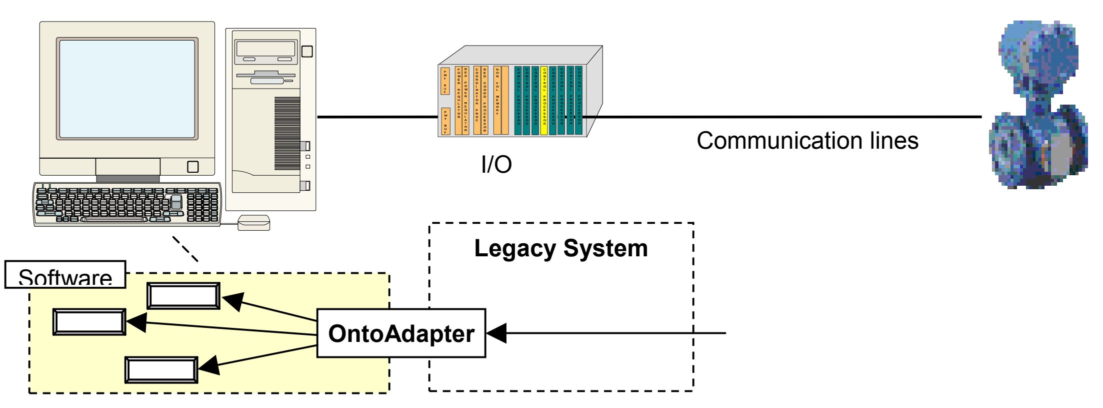

Figure 1. Semantic Adapter is a kind of *"device driver"* for software that uses it

concepts (such as hardware descriptions, hardware diagnostic methods descriptions, etc.). Concrete data will be annotated in terms of upper- and common ontology. Here, ontology provides a basis for a well-understood "common language" to be used between hardware and software.

Adaptation of data sources includes development of configurable semantic wrapper, which will transform native resource data into ontologically annotated, and connector for establishing data link to source of data (see Figure 2). Ontology-based processing tools are future applications, which will be oriented on ontology-based data representation, supporting, for example, provided by W3C Resource Description Framework for it.

3. Visual Interface for Semantic Adapter

The described Semantic Adapter as software must be independent from the ontology that contains metadata for wrapping and from data sources. The participation of a human in a process of binding ontology concept and data source is evident. Only human knows the semantics of the data provided by data sources. Thus, there must be a handy visual interface for binding of data source and metadata notion. The schema of such interface is showed on figure 3.

First of all user must load to the system ontology which will be the base for semantic wrapping of data sources. The loaded ontology can be displayed in the user interface window in form of RDF-graph. The software must use some module, which would allow efficient presentation of graphs on a screen. As one of the tools for displaying of directed graphs *dot* can be mentioned [10]. It is freeware and quite powerful.

After loading the ontology user must select one of the supported by this software data sources from the list or tree. The picture of the selected data source with its attributes is placed on the binding field (visual component). Then user defines data flows from data sources to ontology nodes (dotted arrows). The bindings are saved in software core in some internal format. After this process software (Semantic Adapter) will know what data source corresponds to the requested semantic data.

4. Conclusions

The described idea of Semantic Adapter supports next-generation Semantic Web enabled ICT products. The latter conform much to the growing market of inter-enterprise resource sharing and managing architectures. These market trends seems to be long-term, hence the training of research experts and developers in this domain seems to be very reasonable.

Ontology-based information description opens new possibilities to describe not only data, but also data processing. It is assumed that it will be possible to annotate diagnostic methods, algorithms, schemas in integration environment, so they can be exchanged. Ontologies and Semantic Web ideology, as a comprehensive approach, are the candidate technology for doing. Thus, designing software supporting semantic wrapping of domain-specific data as shown above is one of the first steps in building complex, flexible domain-specific environment.

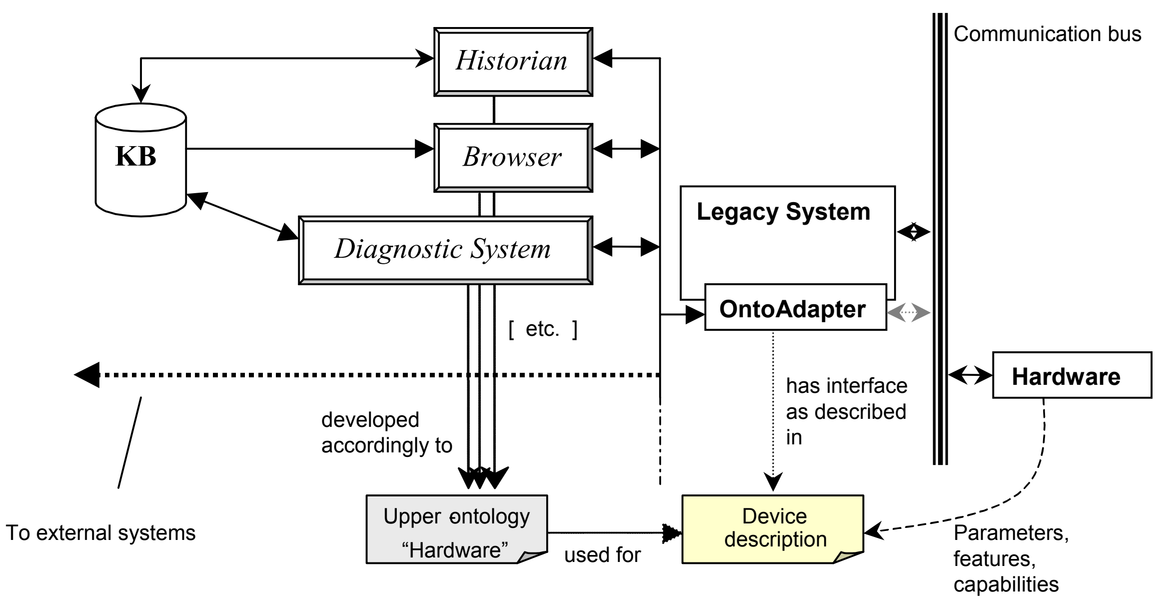

Figure 2. Ontology-based standardization of information exchange between monitored hardware and controlling systems

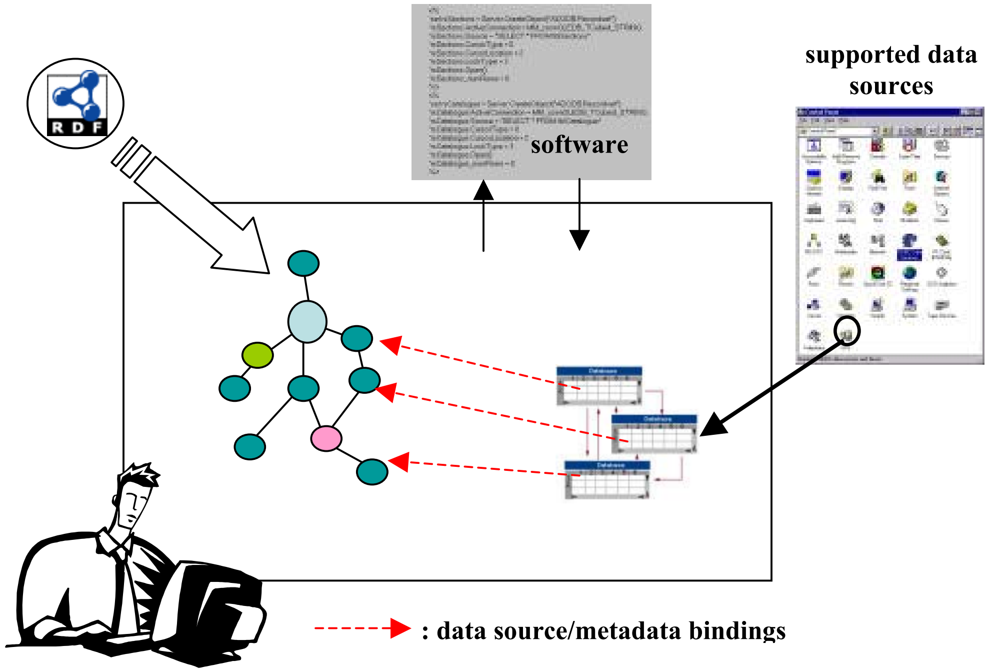

Figure 3. Schema of visual interface for Semantic Adapter configuring

The proposed in this paper idea of visual interface for configuring Semantic Adapter requires further research in domain of user interfaces. The chosen way of configuration will be intuitively understandable for users, who don't have knowledge about Ontology presentation standards.

Actually, Semantic Adapter is one of the constituents of GUN - *Global Understanding eNvironment,* being developed by Industrial Ontologies Group [11]. GUN concept brings extension to Semantic Web vision, making passive semantically annotated resources active. The idea is to make resources capable of being an active part of existing environment, where it resides. This introduces new possibilities to organize heterogeneous resources into dynamic goal-driven systems.

REFERENCES

[1] Basex Inc., "Executive report: Clerity for Enterprise Knowledge Sharing", January 2002.

[2] G. Hohpe, "Enterprise Integration Patterns", ThoughtWorks Inc., 2003.

[3] D. Fensel, F. Harmelen, M. Klein, H. Akkermans, "On-To-Knowledge: Ontology-based Tools for Knowledge Management", Proc. eBusiness and eWork 2000 Conf. (EMMSEC 2000), 2000.

[4] Ph. Perez et. al., "Corporate Memory management Through Agents", Proc. of the eBusiness and eWork Conference, Madrid, Spain, 2000.

[5] M. Ehrig, C. Tempich, S. Staab, F. Harmelen, R. Siebes, M. Sabou, J. Broekstra, H. Stuckenschmidt, "SWAP: Ontology-based Knowledge Management with Peer-to-Peer", 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'03), London, 2003.

[6] U. Thaden, W. Siberski, W. Nejdl, "A Semantic Web based Peer-to-Peer Service Registry Network", Technical Report, 2003.

[7] K. Sivashanmugan, K. Verma, R. Mulye, Z. Zhong: "SpeedR: Semantic Peer-to-Peer Environment for Diverse Web Service Registries", Final Project, CSCI 8350: Enterprise Integration, December 16th, 2002.

[8] http://www.w3.org/ - Official web site of WWW Consortium.

[9] http://www.w3.org/2001/sw/ - Web page of Semantic Web Activity of W3C.

[10] E. Koutsofios, S. North. "Drawing graphs with *dot"*. Technical Report 910904-59113-08TM, AT&T Bell Laboratories, Murray Hill, NJ, September 1991.

[11] http://www.cs.jyu.fi/ai/OntoGroup/ - official website of the Industrial Ontologies Group.