Data Interoperability via Ontology-Based Data Understanding – A Brief Overview

Abstract

Data interoperability is one of the grand challenges when reengineering information systems. Data integration is a difficult task since data source can be heterogeneous in syntax, schema, or semantics. In order to achieve semantic interoperability in a heterogeneous information system, the meaning of the integrated data has to be understood across the systems, which is defined as data understanding in this study. Ontology has been investigated in the context of knowledge representation among heterogeneous and disparate information systems. Our recent study and experiments [38-41] suggest that ontology also has a great potential for heterogeneous data understanding during software reengineering. The main idea is using ontology as data standardisation and conceptualisation via a formal machine-understandable ontology language, which will act as a mediator for reconciliating and understanding the heterogeneities between different data sources during system evolving.

Keywords: Ontology, Software Reengineering, Data Understanding, Data Interoperability, Data Integration

1. Introduction and Motivation

Existing available information is often heterogeneous and distributed in legacy data centric systems, while completely accessing to them is highly demanded in system evolution. Many technical problems have to be solved in the first place in order to establish an efficient information sharing framework when evolving these data centric systems.

Firstly, a suitable relevant data source must be easy to be found as long as it might contain data needed for a given task. Finding suitable information sources is a problem addressed in the areas of information retrieval and information filtering [5].

Once the target data source has been located, the access to the data source has to be provided correspondingly. This means that each of the data sources found in the first step needs to be able to work together with the system that is querying the information. The problem of bringing together heterogeneous and distributed data sources is known as data interoperability problem.

Data interoperability is one of the grand challenges when reengineering data centric information systems. It is relevant to the evolutions of a number of systems including enterprise information systems, medical information systems, geographical information systems, and E-Commerce systems. In general, data source in these systems can be heterogeneous in syntax, schema, or semantics, thus making data interoperation a difficult task [6].

- Using different models or languages will cause syntactic heterogeneity.
- Different structure will lead to schematic heterogeneity.
- Different meanings or interpretations of data in various contexts will bring up semantic heterogeneity.

Here in this study, we mainly focus on semantic heterogeneity. Goh identifies three main causes for semantic heterogeneity [16]:

• **Confounding conflicts** occur when information items seem to have the same meaning, but differ in reality, e.g. owing to different temporal contexts.

• **Scaling conflicts** occur when different reference systems are used to measure a value. Examples are different currencies.

• **Naming conflicts** occur when naming schemes of information differ significantly. A frequent phenomenon is the presence of homonyms and synonyms.

As described in [11], the emergence of XML has created a syntactic platform for Web data standardization and exchange. However, schematic data heterogeneity may still persist, depending on the XML schemas used (e.g., nesting hierarchies). Likewise, semantic heterogeneity may persist even if both syntactic and schematic heterogeneities do not occur (e.g., naming concepts differently).

In order to achieve semantic interoperability in a heterogeneous information system, the meaning of the data source that is to be integrated has to be understood across the systems, which is defined as data understanding in this study. Semantic conflicts occur whenever two contexts do not use the same interpretation of the information.

There are two most important approaches for building a data integration system: Global-as-View (GaV) and Local-as-View (LaV) [25, 35]. In the GaV approach, every entity in the global schema is associated with a view over the data source local schema. Therefore querying strategies are simple, but the evolution of the local data source schemas is not easily supported. On the contrary, the LaV approach permits changes to data source schemas without affecting the global schema, since the local schemas are defined as views over the global schema, but query processing can be complex.

Uschold and Gruninger mention interoperability as a key application of ontology, and many ontology-based approaches [36] to information integration in order to achieve interoperability have been developed. Ontology has been extensively used in data integration systems because they provide an explicit and machine-understandable conceptualisation of a domain. To facilitate data integration, they have been used in one of the three following ways [37]:

• **Single ontology approach.** All data source schemas are directly mapped to a shared global ontology that provides a uniform interface to the user. However, this approach requires that all data sources have nearly the same level of granularity, with the same perspective on a domain. A typical example of a system using this approach is SIMS [1].

• **Multiple ontology approach.** Each data source is represented in its own (local) ontology separately. Local ontologies are mapped to each other instead of mapped to a global one. In addition, additional representation formalism will be needed to define the inter-ontology mappings. The OBSERVER system [27] is an example of this approach.

• **Hybrid ontology approach.** A combination of the two previous approaches is introduced as the third but the most effective approach. Local ontology is built for each data source schema, without mapping to other local ontologies, but to a global ontology. New data sources can be easily added by not modifying existing mappings.

The single and hybrid approaches are appropriate for building central data integration systems, the former being more appropriate for GaV systems and the latter for LaV systems.

In our study, knowledge representation and reengineering techniques will be explored to develop an ontology-based reengineering approach to achieve data interoperability in data centric system evolution.

2. Research Hypothesis

The main hypothesis underlying this study is that ontologising data sources are useful means to understand heterogeneous data and to achieve data interoperability during data centric software system reengineering in a cost efficient manner.

H1: Semantic heterogeneity is the most crucial factor which affects data interoperability.

H2: Neither of Gav and Lav is sufficient to data integration, but combination of these two techniques and knowledge based techniques may overcome the shortcomings of traditional techniques.

H3: There always exists a proper way to extract ontology from existing software system, assuming that existing system may be an object-oriented system or a procedure-based system.

H4: There exist adequate available domain ontology resources for domain-specific software system.

H5: Since only parts of data understanding process can be automated, a cost efficient and highly industrialised data understanding approach will have to support the following features:

- Provide a well defined data understanding methodology.
- Maximise automation, but allow for flexibility and combination with manual methods.
- Provide a decision framework for automation vs. manual activities, based on cost and resulting quality.

H6: Only small parts of existing system are representing real business rules and process logic as domain knowledge.

3. Proposed Solution

We call data understanding the process of using a conceptual representation of the data and of their relationships to eliminate possible heterogeneities and to achieve data interoperability. At the heart of data understanding is the concept of ontology, which is an explicit specification of a shared conceptualisation [17, 18]. Ontologies were developed by the Artificial Intelligence community to facilitate knowledge sharing and reuse [19]. Carrying semantics for particular domains, ontologies are largely used for representing domain knowledge.

3.1. Use of Ontology

An ontology is a system of concepts in which all concepts are defined and interpreted in a declarative way. Such system defines the vocabulary of a problem domain and a set of constraints on the way terms can be combined to model the domain [9, 12]. Description Logic (DL) [4], which is a knowledge representation formalism, can be used as a specification language for a formal ontology. Such usage enables reasoning services provided by DL-based knowledge representation systems. Using this DL-based ontology representation and its inference mechanism will provide a guidance and support during knowledge acquisition in software migration [28]. To retrieve the information from the ontology, the ontology query languages will be needed. The basic theory of such query languages can be divided into two mechanisms: RDF-based query and Logic/Rule-based query. RDF-based query is based on matching RDF triple notation with RDF graph, e.g., SPARQL [32], while Logic/Rule-based query is based on reasoning services provided by logic and rules, e.g., DIG interface [13] and SWRL [21]. Moreover, SPARQL is lack of semantics, while DIG interface and SWRL will be more capable in semantic services and provide proved algorithm to retrieve information that can be inferred from the ontology. Such ontology representation as well as its query language and DL-based inference mechanism will provide a guidance and support during knowledge acquisition in software reengineering process.

Initially, ontology is introduced as an “explicit specification of a conceptualisation” [17]. Therefore, ontology can be used in data understanding and integration task to describe the semantics of the information sources and to make the contents explicit. With respect to the comprehension of data sources, they can be used for the identification and association of semantically corresponding information concepts.

3.2. Generation of Ontology

There exist many approaches to populate ontology from various data sources.

- Existing ontology library
- Text mining technique
- Reverse engineering technique

The proposed ontologies can be divided into two categories: domain ontology and dataset ontology.

To generate domain ontology, existing ontology library can be utilised. In addition, the software documentation will also be explored if available. Techniques such as text mining can be hired to generate domain ontology from software document.

On the other hand, to generate dataset ontology, reverse engineering techniques will be employed. E.g. Reverse engineering technique to support transformation from database schema to ontology will be explored. Since ontology is not only about concepts and instances, relationships are also crucial parts in ontology. Extracting and mapping relationships in a relational database will become very important in terms of ontology generation.

3.3. Use of Mappings

Ontologies should be seen as the glue which sticks various data sources together instead of the standard data format which are transformed from various data format. Consequently, the mapping from an ontology to its environment plays an essential role in information integration [37]. Term mapping is used to refer to the link between ontology and other parts of the application system. In [37], two most important uses of mappings required for information integration have been discussed: mappings between ontologies and the information they describe and mappings between different ontologies used in a system.

Mapping from ontology to the actual contents of data source is the first and most obvious application of mappings in this study. From this aspect, not only ontology could be linked to database scheme, but also it could relate to the single terms in the data source. Wache et al. [37] observe different general approaches used to establish a connection between ontologies and data sources. They briefly discuss these general approaches in their study.

- Structure Resemblance: simply performs a one-to-one copy of the database structure and encodes it into ontology language which provides reasoning services. This approach is implemented in the SIMS mediator [2] and also by the TSIMMIS system [8].
- Definition of Terms: BUSTER [34] uses the ontology to further define the terms or vocabularies in the database or the database scheme. Instead of copying the structure of the database, these vocabularies are only linked to the terms used in data sources.
- Structure Enrichment: combines the two former approaches, and is the most common approach to mapping ontology to data source. A detailed discussion of this kind of mapping is given in [24]. Systems that use structure enrichment for information integration are OBSERVER [27], KRAFT [31], PICSEL [15] and DWQ[7].
- Meta-Annotation: uses meta-annotations to add semantic information to data source. This approach is prominent with the need to integrate data sources stored in the World Wide Web in order to achieve semantic web. Ontobroker [14] and SHOE [20] are examples for this approach.

Mapping between different ontologies is the second crucial part of this study. Many of the existing information integration systems such as [27] or [31] use more than two ontologies to represent data. The problem of mapping different ontologies is a well known problem in knowledge engineering. Following approaches has been surveyed in [37].

- Defined Mappings: In KRAFT [31], mappings are defined in first place, and translations between different ontologies are done by special mediator agents based on the defined mappings.
- Lexical Relations: In the OBSERVER system [27], intuitive semantics for mappings between concepts in different ontologies are provided. The approaches extend a common description logic model by quantified inter-ontology relationships borrowed from linguistics. The subsumption algorithm is rather heuristic than formally grounded.
- Top-Level Grounding: To prevent from losing semantics, one has to stay inside the formal representation language when defining mappings between different ontologies (e.g. DWQ [7]). This could be done by a single top-level ontology, i.e., inheriting concepts from a common top-level ontology. This approach can be used to resolve conflicts and ambiguities (compare [20]).
- Semantic Correspondences: To avoid arbitrary mappings between concepts, these approaches have to rely on a common vocabulary for defining concepts across different ontologies.

Both of these two mappings are crucial in our study. The former one is more related to reengineering techniques while the latter one is more connected to knowledge engineering techniques. Building these two types of mappings will be the key to our approach.

3.4. Ontology Engineering

The previous sections provided information about the use and importance of ontologies. Hence, it is crucial to support the development of ontology.

Development Process

Uschold and Gruninger [36] defined four main phases:

Phase 1: Indentify Purpose and Scope

Phase 2: Building the ontology

(a) Ontology capture: Knowledge 1) identification of the key concepts and relationships; 2) production of precise unambiguous text definitions; 3) identification of terms; agreeing on all of the above.

(b) Ontology coding: Explicit representation of the conceptualisation in some formal language

(c) Integrating existing ontologies: Reuse of existing ontologies to speed up the development.

Phase 3: Evaluation

Phase 4: Guidelines for each phase

Development Principles

A set of development principles will be concluded to guide the design and construction of ontology in our study.

Supporting tools

An open-source platform will be selected to support ontology development, e.g., Protege, etc.

4. Case Study

The Leicester Digital Mapping project is based in the Institute Of Creative Technologies at De Montfort University. The success of the Leicester Digital Mapping project depends upon the range, accuracy and usefulness of the data it incorporates, and upon the ability of large and diverse data sets to overlay and interact with one another. This part of the project therefore focuses upon data gathering and integration, and the creation of an underlying computational infrastructure (DARBS) that is capable of uniting the various elements.

The ontology-based data understanding approaches will be designed to achieve data interoperability in data integration part of this project. Geographical ontology will be used in data understanding and integration for following purposes:

- **Metadata**: Definitions for the basic atomic information in the data source – data about data.
- **Global Vocabulary**: Geographical thesaurus which will be used as reference for terms in all data sources.
- **Data Glue**: Not only stick the semantic heterogeneous graphical data together, but also adhere related data from different aspect. E.g., graphical information, historical information and cultural information
- **Knowledge Based Queries**: Knowledge based reasoning engine will enable precise and efficient query in integrated data sources.

5. Expected Contributions

This data understanding and integration approach is proposed in the context of knowledge representation, which is the application of description logic and ontology to the task of constructing computable models for some domain. Concretely, the expected contributions are as follow:

C1: A series of ontology design principles will be proposed to guide and facilitate data source ontology design.

C2: A great deal of effort, including the definition of basic terms and relations in data source, is devoted to define ontology for data source.

C3: A framework for ontology-based data understanding is presented. Methodology of data source ontology generation and algorithm for data source ontology mapping are designed.

C4: A set of toolsets will be developed to demonstrate the effectiveness of the proposed approach by understanding data source in existing data centric system.

6. Evaluation

A set of metrics will be designed to measure and analyse the proposed study, and therefore to evaluate the efficiency and effectiveness of it.

A whole criterion for the success of proposed study is how well it supports successful automated data understanding and integration. The following criteria are given to judge the success of the research described in this study:

- to be able to deal with as many kinds of existing data centric systems as possible.
- to support the modern data source formats.
- to be feasible for realisation. For instance, it is possible to build a practical tool to demonstrate the approach.
- to be capable for industrial-scaled systems.

7. Related Work

This section will introduce several existing research activities on ontology generation and data integration [37].

Infosleuth: This system semi-automatically constructs ontology from textual databases [22]. The methodology is as follows: human experts provide a small number of seed words to represent high-level concepts; the system then processes the incoming documents, extracting phrases that involve seed words, generates corresponding concept terms, and then classifies them into the ontology; during this process the system also collects seed word-candidates for the next round of processing; as more documents arrive, the ontology expands and the expert is confronted with the new concepts.

KRAFT: offers two methods for building ontologies: the building of shared ontologies [23] and extracting of source ontologies [30].

Shared ontologies: The steps of the development of shared ontologies are:

(1) ontology scoping,

(2) domain analysis,

(3) ontology formalisation,

(4) top-level-ontology.

Extracting ontologies: Pazzaglia and Embury [30] introduce a bottom-up approach to extract an ontology from existing shared ontologies. This extraction process consists of two steps. The first step is a syntactic translation from the KRAFT exportable view (in a native language) of the resource into the KRAFT-schema. The second step is the ontological upgrade, a semi-automatic translation plus knowledge-based enhancement, where local ontology adds knowledge and further relationships between the entities in the translated schema.

Ontobroker: They distinguish between three classes of web information sources (see also [3]) in the first instance:

(1) Multiple-instance sources with the same structure but different contents,

(2) single-instance sources with large amount of data in a structured format, and

(3) loosely structured pages with little or no structure.

Ontobroker [Decker et al., 1999] has two ways of formalizing knowledge (this refers to phase 2b). Firstly, sources from (1) and (2) allow implementing wrappers that automatically extract factual knowledge from these sources. Then sources with little or no knowledge have to be formalized manually.

SIMS: An independent model of each information source must be described for this system, along with a domain model that must be defined to describe objects and actions [1]. SIMS model of the application domain includes a hierarchical terminological knowledge base with nodes representing objects, actions, and states. In addition, it includes indications of all relationships between the nodes. Further, the authors address the scalability and maintenance problems when a new information source is added or the domain knowledge changes. As every information source is independent and modeled separately, the addition of a new source should be relatively straightforward. The domain model would have to be enlarged to accommodate new information sources or simply new knowledge (see also [26]).

To add more...

8. Conclusion

Data integration is a difficult task since data sources can be heterogeneous in syntax, schema, or semantics. Ontology provides an explicit and formal specification of a shared conceptualization, and is able to facilitate knowledge sharing and reuse. Ontology generation and ontology mapping will be two crucial parts in this study. Ontology can be populated manually or (semi-)automatically from various knowledge sources (e.g., database schemas) based on some reengineering techniques. Techniques used for ontology mapping, including ontology alignment and ontology merging [10, 29], overlap to a large extent with those techniques for schema matching [33]. In the future, more details related to the underpinning techniques will be researched and case study based on Leicester Digital Mapping Project will be carried out.

References

[1] Y. Arens, C. Y. Chee, C. N. Hsu and C. A. Knoblock, "Retrieving and Integrating Data from Multiple Information Sources", *International Journal of Cooperative Information Systems,* vol. 2(2), pp. 127-158, 1993.

[2] Y. Arens, C. Knoblock and C.-N. Hsu, "Query Processing in the SIMS Information Mediator", in *Advanced Planning Technology*: AAAI Press, 1996, pp. 61-69.

[3] N. Ashish and C. A. Knoblock, "Semi-automatic Wrapper Generation for Internet Information Sources", *2nd IFCIS International Conference on Cooperative Information Systems (CoopIS)*, Charleston, SC, 1997.

[4] F. Baader, D. Calvanese, D. McGuinness, D. NardiPeter and Patel-Schneider, "The Description Logic Handbook": Cambridge University Press, Jan. 2003.

[5] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: two sides of the same coin?" *Communications of the ACM*, vol. 35(12), pp. 29 - 38, 1992

[6] Y. Bishr, "Overcoming the semantic and other barriers to GIS interoperability", *International Journal of Geographical Information Science*, vol. 12(4), pp. 299 - 314, Jan. 1998.

[7] D. Calvanese, G. D. Giacomo and M. Lenzerini, "Description Logics for Information Integration", in *Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski*: Springer-Verlag, 2001, pp. 41 - 60.
[8] S. Chawathe, H. Garcia-Molina, J. Hammer, K. Ireland, Y. Papakonstantinou, J. Ullman and J. Widom, "The TSIMMIS Project: Integration of Heterogeneous Information Sources", *16th Meeting of the Information Processing Society of Japan*, 1994, pp. 7-18.
[9] F. Chen, H. Guo, L. Dai and H. Yang, "An Application Framework for Ontology-based Data Mining", *Journal of Dalian University of Technology*, vol. Suppl., 43(S1), China, pp. 143-145, Oct. 2003.
[10] I. F. Cruz, W. Sunna and A. Chaudhry, "Semi-Automatic Ontology Alignment for Geospatial Data Integration", in *Geographic Information Science*, 2004, pp. 51-66.
[11] I. F. Cruz and H. Xiao, "The Role of Ontologies in Data Integration", *Jornal of Engineering Intelligent Systems*, vol. 13(4), Dec. 2005.
[12] V. Devedzic, "Understanding Ontological Engineering", *Communications of the ACM*, vol. 45(4), pp. 136-144, Apr. 2002.
[13] DIG, "DIG Interface": http://dig.sourceforge.net/.
[14] D. Fensel, S. Decker, M. Erdmann and R. Studer, "Ontobroker: The Very High Idea", *11th International Flairs Conference (FLAIRS-98)*, Sanibal Island, Florida, 1998.
[15] F. Goasdoue, V. Lattes and M.-C. Rousset, "The use of CARIN language and algorithms for Information Integration: the PICSEL project", *International Journal on Cooperative Information Systems*, vol. 9(4), pp. 383-401, 1999.
[16] C. H. Goh, "Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Systems". vol. PhD: Massachusetts Institute of Technology, 1997.
[17] T. R. Gruber, "A translation approach to portable ontology specifications." *Knowledge Acquisition*, vol. 5(2), pp. 199 - 220, 1993
[18] T. R. Gruber and G. R. Olsen, "An Ontology for Engineering Mathematics", *4th International Conference on Principles of Knowledge Representation and Reasoning*, 1994, pp. 258-269.
[19] N. Guarino, "Formal Ontology and Information Systems", *1st International Conference on Formal Ontologies in Information Systems(FOIS'98)*, Trento, Italy, Jun. 1998, pp. 3-15.
[20] J. Heflin and J. Hendler, "Semantic Interoperability on the Web", *Extreme Markup Languages 2000*, 2000.
[21] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof and M. Dean, "SWRL: A Semantic Web Rule Language": http://www.w3.org/Submission/SWRL/, May 2004.
[22] C. H. Hwang, "Incompletely and Imprecisely Speaking: Using Dynamic Ontologies for Representing and Retrieving Information", *6th International Workshop on Knowledge Representation meets Databases (KRDB'99)*, Jul. 1999, pp. 14-20.
[23] D. Jones, "Developing Shared Ontologies in Multi-agent Systems", *ECAI'98 Workshop on Intelligent Information Integration*, Brighton, U.K, 1998.
[24] V. Kashyap and A. Sheth, "Schematic and Semantic Similarities between Database Objects: A Context-based Approach", *International Journal on Very Large Data Bases (VLDB)*, vol. 5(4), pp. 276-304, 1996.
[25] M. Lenzerini, "Data Integration: A Theoretical Perspective." *21st ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS 2002)*, Madison, Wisconsin, Jun. 2002, pp. 233 - 246.
[26] R. MacGregor, "The Evolving Technology of Classification-Based Knowledge Representation Systems", in *Principles of Semantic Networks: Explorations in the Representation of Knowledge*, J. Sowa, Ed., 1990.
[27] E. Mena, A. Illarramendi, V. Kashyap and A. P. Sheth, "OBSERVER: An Approach for Query Processing in Global Information Systems Based on Interoperation Across Pre-Existing Ontologies", *Distributed and Parallel Databases*, vol. 8(2), pp. 223-271, Apr. 2000.
[28] W. J. Meng, J. Rilling, Y. Zhang, R. Witte, S. Mudur and P. Charland, "A Context-Driven Software Comprehension Process Model", *2nd International IEEE Workshop on Software Evolvability (SE'06)*, Sep. 2006, pp. 50-57.
[29] N. F. Noy and M. A. Musen, "PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment ", *17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence*, 2000, pp. 450-455.
[30] J.-C. R. Pazzaglia and S. M. Embury, "Bottom-up Integration of Ontologies in a Database Context", *5th INternational Workshop on Knowledge Represenation Meets Databases (KRDB '98)*, Washington, USA, 1998.

[31] A. Preece, K. Hui, A. Gray, P. Marti, T. J. M. Bench-Capon, D. M. Jones and Z. Cui, "The KRAFT Architecture for Knowledge Fusion and Transformation", *19th SGES Int. Conf. on Knowledge-based Systems and Applied Articial Intelligence (ES'99)*, 1999.
[32] E. Prud'hommeaux and A. Seaborne, "SPARQL Query Language for RDF: W3C Working Draft 26 March 2007", http://www.w3.org/TR/rdf-sparql-query/, Mar. 2007.
[33] E. Rahm and P. A. Bernstein, "A Survey of Approaches to Automatic Schema Matching", *International Journal on Very Large Data Bases*, vol. 10(4), pp. 334-350, 2001.
[34] H. Stuckenschmidt and H. Wache, "Context Modelling and Transformation for Semantic Interoperability", *Knowledge Representation Meets Databases (KRDB 2000)*, 2000.
[35] J. D. Ullman, "Information Integration Using Logical Views", *6th International Conference on Database Theory (ICDT 1997)*, 1997 pp. 19 - 40.
[36] M. Uschold and M. Grüninger, "Ontologies: Principles, methods and applications", *Knowledge Engineering Review*, vol. 11(2), pp. 93-155, 1996.
[37] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann and S. Hübner, "Ontology-Based Integration of Information A Survey of Existing Approaches", *the International Workshop on Ontologies and Information Sharing*, 2001, pp. 108-117.
[38] H. Yang, Z. Cui and P. O'Brien, "Extracting Ontologies from Legacy Systems for Understanding and Re-Engineering", *23rd Annual International Computer Software and Applications Conference (COMPSAC'99)*, Phoenix, AZ, Oct. 1999, pp. 21-26.
[39] H. Zhou, F. Chen and H. Yang, "Developing Application Specific Ontology for Program Comprehension by Combining Domain Ontology with Code Ontology", *8th International Conference on Quality Software (QSIC'08)*, Oxford, UK, Aug. 2008.
[40] H. Zhou, J. Kang, F. Chen and H. Yang, "OPTIMA: an Ontology-based PlaTform-specIfic software Migration Approach", *7th International Conference on Quality Software (QSIC'07)*, Portland, Oregon, USA, Oct. 2007, pp. 143-152.
[41] H. Zhou, "COSS: Comprehension by Ontologising Software System", *24th IEEE International Conference on Software Maintenance (ICSM'08)*, Beijing, China, Sep. 2008.