Semantic Integration of Knowledge Management Systems

Napoleon - Alexandru SIRETEANU, "Alexandru Ioan Cuza" University, Iaşi, Romania, salexis@feaa.uaic.ro
Constantin Florin Sîrbu, "Alexandru Ioan Cuza" University, Iaşi, Romania, florinsarbu@feaa.uaic.ro

Abstract

Ontologies are essential for a semantic integration of knowledge components which can be seen as superior pieces of intellectual capital. Knowledge structures connect individual or group knowledge elements from organizational knowledge bases that have been developed independently. This can aid search and navigation for the entire organizational knowledge base. As many knowledge processes cross boundaries of organizational units or of whole organizations, standardization of techniques for representing knowledge structures (Semantic Web technologies), play an important role in enabling organizations to share documented knowledge.

1. Introduction

Whatever the nuances of its definitions used in the management information systems, knowledge is extracted from an expert and made available in particularly designed systems and knowledge-based systems. Alternatively, management is a term that denotes the software-supported handling, e.g., storing, administering, updating and retrieving of business objects when used in connection with information and communication technology (ICT). Examples are database management systems or document management systems. The term KMS (knowledge management system) has been a strong symbol or vision for the development of a new type of ICT systems.

Knowledge management has had an upstream evolution from the age of data and information to that of the economy of knowledge, as depicted in Table 1. The table also suggests the way each age has treated its corresponding management topics and technological elements [5].

Table 1: Suitability of the knowledge age

Evolution				
Management Concepts	Systemic thinking/approach System/project management	Software engineering management CMM	System engineering management CMM	Knowledge engineering management KMA/EE
Systems	Data Processing Systems (DPS)	Management information systems (MIS)	Decision support systems (DSS)	Knowledge management systems (KMS)
Technological Elements	Data	Information	Artificial intelligence	Knowledge
Age	*Industrial*	*Technology*	*Information*	*Knowledge*

2. Knowledge management systems

Knowledge management systems create a corporate ICT environment, a contextualized base, an infrastructure that takes into account the complex nature of knowledge and supports the handling of knowledge in organizations. In order to achieve this, a number of heterogeneous ICT have to be integrated, improved, recombined and repackaged. Examples are AI technologies, business intelligence technologies, communication systems, content and document management systems, group support systems, learning environments, search engines and workflow management systems.

ICT systems have been developed to provide sophisticated functions for publication, organization, visualization, contextualization, search, retrieval and distribution of knowledge as well as functions supporting communication, collaboration, cooperation and linking of individuals in social networks. Many organizations experiment a set of technologies known as Web 2.0 or social software. Examples are forums, Wikis, Weblogs, "social" bookmarking or tagging solutions.

The growing speed of innovation in the field of ICT has provided numerous technologies ready to be applied in organizations. Examples of information and communication technologies that are related to knowledge management are:

- *Intranet infrastructures* provide basic functionality for communication (e.g., e-mail), as well as storing, search and retrieval of data and documents;
- *document and content management systems* manage electronic documents or Web content during their entire life cycle. A document management system (DMS) provides functions to store and archive documents, navigate and search documents, manage versions and to control access to documents. Like DMS in the

non-Web environment, CMS (content management systems) manage the whole Web publishing process, offer mechanisms for releasing new contents, support HTML generation by using templates, standard input and output screens and the separation of content and layout which provides for a standardized aspect of the Web pages.

Content management systems allow separating the editorial activities from the layout, facilitating the maintenance of an Internet presence. This means, however, that the structure of an Internet presentation has to be mapped. The specifics of content management systems are that special tools for various participating roles (e.g., graphic artists or editors), are available, while HTML knowledge is normally not required. Another benefit is that content, layout, and navigation are separate, the contents of single information units are specified, and workflows can be mapped. Wikis and Weblogs use specialized CMS that are pre-structured, offer an easy to use CMS functions and allow editing, updating and linking of content within and between sites [6];

- *workflow management systems* support well-structured organizational processes, handle the execution of workflows, and consists of activities related to one another which are triggered by external events and carried out by persons using resources such as documents, application software and data;
- *artificial intelligence technologies* support search and retrieval, user profiling, text and Web mining;
- *data warehouse* is an integrated and non-volatile collection of data in support of management decision processes and is physically separated from operational systems;
- *business intelligence tools* support the analytic process which transforms fragmented organizational and competitive data into “knowledge” and require an integrated data that is frequently provided by a data warehouse. Technologies that support this process are decision support system (DSS) technologies, multidimensional analysis (OLAP), data mining, text mining and Web mining technologies, business simulation techniques, and artificial intelligence technologies;
- *groupware and collaboration software* supports for meetings or creative workshops of work groups and teams. Groupware tools can include communication systems (email, audio/video systems, chat systems), information sharing systems (message boards), cooperation systems (co-authoring, word processor, spreadsheet, group decision support systems), co-ordination systems (notification systems) and social systems (media spaces, virtual reality) [2];
- *e-learning systems* offer specified learning content to employees in an interactive way and thus support the teaching and/or learning process.

More sophisticated tools, such as knowledge databases that archive unstructured knowledge resources in ways that can quickly be found through keyword searches, form the next step in integrating modern training and knowledge management. E-learning developers work out ways to store and manage learning content in modular, object-based formats. In this way, learning content could be served to users together with other knowledge resources from the same knowledge repository [1].

3. Semantic integration of knowledge management systems

A knowledge management system (KMS) is an ICT system or an ICT platform that combines and integrates functions for the contextualized usage of explicit and tacit knowledge, all over the organization or that part of the organization that is targeted by a KM initiative. A KMS offers integrated services to deploy KM instruments for networks of participants (knowledge workers), in knowledge-intensive business processes along the entire knowledge life cycle. A critical goal of KMS is to support the dynamics of organizational learning and organizational efficiency.

Knowledge management systems might also be viewed as important organizational assets that provide core competencies for the organization. Especially highly knowledge-intensive organizations might view the systematic handling of knowledge in general and their ICT systems supporting KM in particular as their core competence and fear that they might loose a strategic advantage if they implement a standard software solution available on the market.

Integration of data has been a trouble for many years. Relational database management has unified the way (transactional) data is handled in organizations. The amount of semi-structured and unstructured data, such as documents, messages, images, media files or Web content has grown considerably and needs to be integrated as well. The integration of these data sources requires in addition to data integration, semantic integration for standards and technologies. Many of these standards and technologies are built on XML.

Table 2 exemplifies the various developed technologies that made the concept of the Semantic Web possible. As we can see, the Web was initially a vast set of static Web pages (static HTML files) linked together, that are still use by many organizations. Due to the dynamic nature of businesses, organizations are using dynamic publishing methods which offer great advantages over Web sites constructed from static HTML pages. Server-side applications and database access techniques are used to dynamically create Web pages directly in response to requests from user browsers. This technique offers the opportunity to deliver Web content that is highly customized to the needs of individual users. The technologies available to dynamically create Web pages based on database information were insufficient for the requirements of organizations looking for application integration solutions. The Extensible Markup Language (XML) was one of most successful solutions developed to provide business-to-business integration. XML-based solutions for applications and systems integration were not sufficient, because the data exchanged lacked an explicit description of its meaning. Semantic integration and interoperability is concerned with the use of explicit semantic descriptions to facilitate integration [7].

Table 2: Evolution of Web technologies

Technologies/ Characteristics	**Static**	**Dynamic**	**Syntax**	**Semantic**
Coding	HTML	RDMS	XML	RDF/OWL/OIL
Creation	Manually or using web development application	By server-side applications	By application based on schema	By applications based on models
Users	Humans	Humans	Humans and applications	Humans and applications
Paradigm	Browse	Create/Update/Delete queries	Integrate	Interoperate
Applications	Browsers	Browsers	Process Integration, EAI, Workflows	Intelligent agents, Semantic engines

1995 2000 2006

Semantic integration of semi-structured data is a complex task. Semantic Web breaks down the variety of tasks into a layered structure that helps to understand what concepts have to be defined so that semantic information about knowledge elements can be easily exchanged between a diversity of heterogeneous ICT systems.

Semantic integration intends to providing standards for describing documents or, more generally, Web resources. This is done with the help of statements that describe the resources (RDF), a vocabulary for the definition of constraints on these statements (RDF Schema), ontologies that show the relationships between the concepts used in descriptions and vocabularies (OWL) and a logic framework that allows interpretation of documents and descriptions.

Resource Description Framework (RDF) standard is an XML-based language for representing meta-data about Web resources and to connect Web resources to each other. In this model, every type of information about a resource, which may be a web page or an XML element, is indicated in terms of a triple [4]: resource, property, value. RDF provides mechanisms to add semantics to a resource as an individual entity without assumptions about its internal structure. RDF is based on the idea of identifying things using URIs and describing them with properties and property values. An RDF statement consists of a triple of subject, predicate and object.

The RDF Vocabulary Description Language (RDF Schema) supports designing vocabularies (classes for instance RDF specifications). Schemas are needed for describing terms used in RDF statements (types of things, properties). RDF Schema proposes well-defined rules for writing these definitions which can be exchanged and parsed automatically to extract semantics of RDF statements about Web resources.

The Web Ontology Language (OWL) is a language for defining and instantiating Web ontologies that include descriptions of classes, properties and their instances. Compared to RDF Schema, OWL offers more facilities for describing classes and properties, e.g., relations between classes, cardinality, equality, better typing of properties and characteristics of properties. OWL provides a number of language elements that specifically target the integration of concepts defined in different ontologies. OWL is the most powerful of the ontology languages currently defined for the Semantic Web. Unlike DAML+OIL, OWL is sponsored by World Wide Web Consortium (W3C)[3]. OWL has facilities for expressing meaning and semantics and the ability to represent machine interpretable content on the Web. OWL is designed for use by applications that need to process the content of information instead of just presenting information to humans.

While RDF, RDF Schema and OWL have been standardized for quite some time and there are a number of projects in many organizations that are based on these standards, there is still a lot of contest at the higher levels of the Semantic Web stack. Standards that allow for specifying entire logic frameworks, exchanging proofs and building trust between agents still remain evident.

4. Discussions and Conclusions

Data, information, and knowledge exchange is becoming the key issue in contemporary computer technology. Ontologies provide a common understanding of a domain that can be communicated between people and different application systems and will play a major role in supporting information exchange processes in various areas. The use of a single ontology for all application will never be possible. An ontology will never be convenient for all subjects and domains or for a large and varied community such as the Web community.

5. References

[1] Andone, I., Sireteanu, N.A., *Strategies for Technology-Based Learning in Higher Education*, available at Social Science Research Network (SSRN), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1078829, november 2007

[2] Andriessen, J.H.E., *Working with Groupware. Understanding and Evaluating Collaboration Technology*, London, 2003, p.12

[3] OWL Web Ontology Language, http://www.w3.org/TR/owl-features/, retrieved april 2, 2008

[4] Resource Description Framework (RDF) , http://www.w3.org/RDF/, retrieved march 28, 2008

[5] Sireteanu, N.A., Bedrule-Grigoruţă, M.V., *Perspectives of Knowledge Management in Universities*, available at Social Science Research Network (SSRN), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1029990, november 2007

[6] Sireteanu, N.A., *Web Applications Development*, course notes, april 2008

[7] Sireteanu, N.A., *Web Ontology Languages and Semantic Web Services*, available at Social Science Research Network (SSRN), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1079565, december 2007