Service-Oriented Development of Content and Knowledge Provision Tool

Sonja D. Radenkovic, Jelena Jovanovic, Vladan Devedzic
FON - School of Business Administration,
University of Belgrade
Belgrade, Serbia
emails: {sonjafon, jeljov, devedzic}@gmail.com

Abstract—This paper presents the service-oriented development of Content and Knowledge Provision tool, one of the core services developed within IntelLEO, an FP7 project in the area of technology-enhanced learning. The project aims at enhancing cross-organizational Learning and Knowledge Building practices at the workplace. Content and Knowledge Provision tool enables employees to upload different kinds of learning resources into a knowledge repository, annotate them, and (re-)discover relevant learning resources by performing semantic search over the knowledge repository. Hence, this service effectively serves as a content management system and semantic search engine within the IntelLEO.

Keywords—Service-Oriented Architecture, Semantic Web, knowledge management, workplace learning.

I. INTRODUCTION

Semantic Web technologies offer a new approach for information and knowledge management that is largely based on the creation and use of semantics-rich metadata. The main idea of Semantic Web is to create a layer of machine processable data on top of the existing Web in order to enable advanced, automated processing and use of Web content [1]. The primary motivation for applying these novel technologies in existing software systems is to facilitate semantic interoperability and/or integration of software products made by different vendors with existing systems. Thus, Semantic Web technologies complement and further improve the capacity of Service Oriented Architecture (SOA) to enable flexible linking of resources related to traditional architectures and, encourage reusability of these resources. In particular, Semantic Web technologies facilitate the implementation of the SOA concept by enabling semantics-driven identification and integration of required services, and sharing of data between them [3].

One system that implements the SOA concept through the use of Semantic Web technologies has been developed within the Intelligent Learning Extended Organization (IntelLEO) project [4]. This project investigated the IntelLEO paradigm—a learning community emerging as a temporal integration of two or more different business and educational communities characterized by different organizational cultures (industrial, research, and educational) [5]. In fact, an IntelLEO is a kind of extended organization focused on cross-organizational learning and knowledge building. Accordingly, the main goal of the project was to enhance Learning and Knowledge Building (LKB) practices in an IntelLEO. The software framework developed in the scope of the IntelLEO project encourages LKB activities, and makes it easier for learners to initiate and/or take part in these activities within an IntelLEO. The framework is composed of a set of services that interoperate through the common ontology framework. One of these services is the Content and Knowledge Provision (CKP) tool that makes use of Semantic Web technologies and the SOA concept to enable employees to upload different kinds of learning resources into a knowledge repository, annotate them, and (re-)discover relevant learning resources by performing semantic search over the knowledge repository. CKP tool enables employees to allocate learning activities in the form of unstructured documents and knowledge sources in order to acquire some new knowledge required for performing their workplace activities. On the other hand, this tool helps in efficient and effective communication and exchange of information within an extended organization, enabling the management and exchange of learning resources between employees during their learning and work-related practices.

This paper presents the SOA-development of the CKP tool and shows how it allows employees to effectively use knowledge and content from the extended organization’s knowledge/content repositories as well as from all over the Web, without wandering and wasting their time on irrelevant resources. The rest of the paper is organized as follows: the next section presents the IntelLEO framework as a SOA-based system. Section 3 describes the CKP tool layered architecture, whereas Section 4 shows the CKP tool evaluation. Section 5 presents related work, whereas the last section gives the conclusions and indicates directions for future work.

II. INTELEO AS A SOA-BASED FRAMEWORK

The IntelLEO framework (Fig. 1) comprises several core services that can be thought of as abstractions for grouping the framework’s principle functionalities—each core service consists of a number of functionality-specific services. The framework also integrates a set of interlinked ontologies—the IntelLEO ontologies—that allows for unified knowledge...
representation within an extended organization (i.e., an IntelLEO) and semantic interoperability of services.

The data layer of the IntelLEO framework (the bottom most layer on Fig. 1) consists of one or more data repositories (typically, one repository for each organization that participates in an IntelLEO) storing data relevant for the IntelLEO services in a format compliant with the IntelLEO ontologies[23]. The use of ontologies allows for representing and storing data together with their semantics, so the meaning of each piece of data is unambiguously defined, and therefore, data can be more easily shared and (re-)used by different services.

![Figure 1. IntelLEO Framework](image)

The core services of the IntelLEO framework can be divided into two main groups. The first group – Learning and Knowledge Building Services – is composed of the following services: 1) Services for Collaborative Learning (green boxes in Fig. 1), in particular Human Resource Discovery (HRD), Working Group Composition (WGC) and User Monitoring (UM); and 2) Content/Knowledge Provision (CKP). The second group – Harmonization Services (orange boxes in Fig. 1) – includes: 1) Learning Path Creator (LPC); and 2) Organizational Policy (OP) Service. These services are often orchestrated from applications to interact and interoperate with each other. In this paper we focus on the Content and Knowledge Provision functionality of the IntelLEO framework implemented in the CKP tool. Description of all other components of the IntelLEO framework can be found in the project deliverables, specifically [2] and [6], available at the project's Website.

III. CKP TOOL LAYERED ARCHITECTURE

Like the overall IntelLEO framework, the CKP tool is also organized in a SOA manner, with its own specific services forming a layered architecture shown on Figure 2.

A. Persistence Layer

The CKP tool uses the persistence layer of the overall IntelLEO framework (Fig. 1). In fact, this layer is used by all the IntelLEO core services. It provides generic services for storage and retrieval of data originating from LKB activities in an extended organization, as well as data about various kinds of knowledge and learning resources that were used in or resulted from those activities. The data are represented and stored as Resource Definition Framework (RDF) triples compliant with the IntelLEO ontologies. In other words, the IntelLEO ontologies serve as models for storing data in the RDF [24] repositories of the Persistence Layer. This means that data are stored with explicitly defined meaning, and, also, that semantics of connections among data items are made explicit (through ontologies). This further implies that the semantics of the data are directly available for processing by any software component accessing those data through the services of the Persistence Layer.

The services of this Layer hide all the specificities of working with ontologies, RDF, and other related technologies of the Semantic Web stack. So, when needing to access or store data in the repository, the CKP tool (or any other software component) simply works with 'regular' Java classes and interfaces and all the tasks related to the storage, retrieval and update of ontology instance data are handled by the services of the Persistence Layer.

In the following section, we present how the CKP tool, in particular, its Service Layer, makes use of the semantics-rich data of the Persistence Layer to provide users with advanced content/knowledge management and sharing functionalities.

B. Service Layer

Services of this layer allow for personalized content/knowledge retrieval. This means that these services provide users with learning/knowledge resources compliant with the particularities of their learning context (e.g., their present learning goals and competences). These services include: Semantic Annotation service, Tagging service, Service for computing relevance of a learning resource, and Retrieval service.

Semantic annotation service (Fig. 2) provides automatic semantic annotation of learning resources by making use of the annotation services of the KIM platform [8].

The main challenge for using KIM platform in the CKP tool was the extending of PROTON [25] ontology, an upper-level ontology within KIM, with the concepts and the relationships from the , domain ontology - an ontology that formally specifies a specific subject domain, used in the CKP tool. Semantic annotations obtained from this process were clearly specified, easy to understand, and served as a basis for useful applications in the CKP tool.
In other words, the main task of the Semantic Annotation service was to retrieve semantic annotations gathered by the KIM platform, to add these annotations to the specific learning resource as well as to transfer them to the user interface of the CKP tool. Tagging service (Fig. 2) helps to add additional information to a specific learning resource in the form of tags or keywords. The process of adding tags is important both for users who perform the tagging (reflection on the content in order to find the terms that best describe it), and members of the extended organization who might want to use specific learning resources in the future (easier search and discovery).

Service for computing relevance of a learning resource (Fig. 2) computes semantic similarity between the given learning resource and a specific learning goal. In particular, the relevance is computed as semantic similarity between a resource and a specific learning goal.

Semantic similarity is computed using information retrieval techniques, namely TF-IDF [26] and Cosine Similarity [27]. This further means that semantic similarity between a learning resource and a learning goal is calculated by measuring similarity between the term vector found in the learning resource and the term vector of the considered learning goal. Term vector is used to represent both learning resource and learning goal as a vector of identifiers [9]. The concepts used for creating these vectors are obtained through semantic annotation of both the learning resource and the learning goal. Each concept forming a vector of the learning resource is associated with its frequency (i.e., number of occurrences) in that learning resource. On the other hand, a learning goal is composed of competences, and these competences can be complex, i.e., composed of sub-competences. Let us explain this through an example. Suppose we have learning goal LG1 that is composed of competences C1 (annotated with concepts T1 and T2) and C2 (annotated with concepts T3 and T4). Competence C1 is composed of sub-competences C1.1 (annotated with the concept T5) and C1.2 (annotated with concepts T6 and T7), while the competence C1.1 is composed of sub-competences C1.1.1 (annotated with concept T8) and C1.1.2 (annotated with concepts T9 and T10)
concept T9). Presented in the form of a tree, the LG1 learning goal and the associated competences look as follows:

\[
\text{LG1}
\]

\[
\bullet \text{C1(T1, T2)}
\]

\[
\begin{align*}
\circ \text{C1.1 (T5)} \\
& \bullet \text{C1.1.1 (T8)} \\
& \bullet \text{C1.1.2 (T9)} \\
\circ \text{C1.2 (T6, T7)} \\
\bullet \text{C2 (T3, T4)}
\end{align*}
\]

Values associated with concepts T1 – T9 are calculated having in mind the distance between the competence a concept is associated with and the learning goal. In particular, the following simple formula is used: \( V_{Tn} = 1/k \), where \( T_n \) represents the concept \((T_n=1,9)\), \( V_{Tn} \) is the value for that concept, and \( k \) is the distance between the competence the concept \( T_n \) is assigned to and the learning goal (LG1). Accordingly, the values for concepts T1 – T9 are as follows:

- \( V(T1) = 1 \)
- \( V(T2) = 1 \)
- \( V(T3) = 1 \)
- \( V(T4) = 1 \)
- \( V(T5) = 0.5 \)
- \( V(T6) = 0.5 \)
- \( V(T7) = 0.5 \)
- \( V(T8) = 0.33 \)
- \( V(T9) = 0.33 \)

Finally, the semantic similarity between the learning resource \( d_1 \) and the specific learning goal \( d_2 \) is calculated by multiplying the vector of concepts of learning resource \( v(d_1) \) and the vector of concepts of specific learning goal \( v(d_2) \) as follows:

\[
\text{sim} (d_1, d_2) = v(d_1) \cdot v(d_2) \quad (1)
\]

The output is a number between 0 and 1 and it presents the relevance of learning resource for specific learning goal. In addition, this relevance is presented in the form of star-scale in user interface (Figure 6A) indicating if the learning resource is relevant for a specific learning goal.

Retrieval service (Fig. 2) enables seamless retrieval of stored learning resources based on the input that can be a domain-specific concept or tag(s). This service queries the repository of learning resources looking for learning resources that are annotated with domain concept(s) or tag(s) given in the user’s request. The result is a list of the ranked search results. If none of the available resources directly matches the user’s request, this service identifies semantically related domain concepts or tags, finds resources annotated with them and suggests those as potentially useful resources. In order to find similar domain concepts, the service looks for concepts that are more general or more specific to those given in the user’s request. In particular, it makes use of skos:narrower and skos:broader relations for structuring the domain concepts in concept hierarchies. These relations are defined by the SKOS (Simple Knowledge Organization System) ontology [10]. It defines classes and properties for modelling specific subject domains in the form of thesauri, taxonomy, or classification scheme.

To rank the retrieved learning resources, this service makes use of semantic similarity between each of the retrieved learning resources and the user’s profile. Semantic similarity is calculated as Cosine similarity between the vector of a specific resource and user’s profile vector. The vector of a resource comprises all tags and the semantic annotations (concepts from the domain-specific ontology for the learning resources. The user’s profile vector comprises all domain concepts and tags related to the user, his personal learning goals, his competences and learning paths he is following. In addition, it contains concepts and tags that reflect the user’s personal priorities, including general interests, learning history, and acquired competences. Each personal priority contains a weighting factor that affects the value of domain concepts and tags associated with that priority. If the search was done upon request of some other service (e.g., some of the IntelLEO services), similarity is computed between the retrieved resources and some other kind of learning asset (e.g., competence, learning activity, learning path), as requested by the service on whose behalf the search was performed.

C. Application Layer

Application layer is the connection to the “external world” and comprehends Web interface and application logic that uses the functionality of the Service Layer (Fig. 2). This layer implements three types of functionalities gathered from the application cases involved in the IntelLEO project: annotation of learning resources, management of learning resources, and semantic search of a repository of learning resources (Fig. 2). These functionalities are offered through a Web-based interface (see the next section).

D. Implementation

Regarding the actual implementation, the CKP Tool, as well as whole IntelLEO solution is implemented in the Java programming language, to be independent from the underlying platform, and is based upon a number of open source frameworks. Specifically, the Service Component Architecture (SCA) framework Apache Tuscany [15] was used to facilitate the implementation of loosely coupled Core Services. In terms of user interfaces, to achieve cross-browser compatibility, Apache Wicket [16], a Java-based web framework, was used in combination with
Javascript framework JQuery [17]. This approach allows for the IntelLEO platform to be installed and run on different target platforms, and enables users to access the IntelLEO platform from different operating systems and with the majority of common web browsers, as was proved during the evaluation phase.

IV. EVALUATION

The CKP tool was evaluated in an empirical study that lasted two months and included three application cases of the IntelLEO project. The first case (AC1) was about an IntelLEO comprising a big multinational corporation in the automotive sector, a research institute and a university. The second application case (AC2) involved an IntelLEO formed by an SME providing IT services in the e-Engineering and e-Manufacturing sectors, and a university-based research group. The third Application Case (AC3) was about an IntelLEO focused on teacher training; the participating organisations were a Teacher Association and a university. The objective of the evaluation study was to collect feedback from end-users, in the three application cases, concerning the usability and usefulness of the overall IntelLEO framework and its individual components as well as to test to IntelLEO hypotheses, which suggested that a synergy of collaboration and harmonisation services increases the individual motivation for LKB activities, a pre-requisite of organizational responsiveness [7].

The evaluation details are presented in [11]. In the text bellow we present a short overview related to evaluation procedure of the CKP tool, compliance with the scope of this paper.

The study was organized through a series of tasks, set in a specific learning scenario, that the participants had to complete. The tasks were the same in all three application cases to allow for comparison of the results obtained in heterogeneous settings. The study was conducted with participants from AC1, AC2, and AC3. Most of them had university degree (83.3%). In terms of occupations, 31% of participants are teachers, 8% are researchers, 23% are students, 15% are technical employees, 17% are engineers and 6% are categorized as “others”.

They interacted with the services of the IntelLEO framework in five tasks. The forth task was related to the CKP tool; hence, we present the part of the study procedure and the results only for that task. At the beginning of the study session, the participants were familiarized with the learning scenario that was adapted to the particularities of each application case. The fourth task in the learning scenario was to share a learning/knowledge resource by making use of the bookmarking/annotation features of the CKP tool. To complete this task, the participants were asked to navigate to the given URL, initiate the CKP tool, and bookmark/annotate the corresponding Web page (e.g., by adding some tags and/or selecting some of the automatically generated tags, and/or choosing related learning goal(s)). After completing the task, the participants were asked to fill in the corresponding questionnaire. For each feature of the CKP tool, the questionnaire presented the participants with the corresponding screenshot and a question statement, asking them about the perceived usefulness of the tool’s presented at that screenshot. Answers were provided in the form of a 5-point Likert scale (5 – strongly agree; 1 – strongly disagree). An example for a statement would be “When I want to plan my personal learning goals, it is useful to tag an online resource with my personal learning goal.” The objective of this investigation was to find out how useful and relevant are the developed services and functionalities for end-users in the three ACs, and also to examine how useful and relevant the CKP tool and its functionalities are in performing this task, especially w.r.t. the motivational and pedagogical challenges of learning in the workplace.

Evaluation of the CKP tool has shown the importance of CKP tool functionalities for learning and knowledge building (LKB) activities. The users have identified several benefits of the CKP tool for workplace learning. Respondents highlighted the fact that colleagues collaboratively create one repository of learning content, which is annotated and updated bottom-up, but accessible and useable by the whole organization. The requirement to structure and document one’s work-relevant knowledge has been highlighted by users of all three ACs, too. This documentation serves 1) to have one’s own knowledge available at a later stage, and 2) to profit from the reciprocal knowledge exchange between colleagues to split up the burden of documenting important lessons learned among several colleagues. They indicated that the links between knowledge and content must be fully exploited by using a shared repository of learning resources.

V. RELATED WORK

To support the learning and evolving the knowledge of employees at the workplace, the CKP tool focuses on providing distributed services for the semantic annotation, managing and semantic search of learning resources in extended organizations. Many tools/services have supported knowledge management at the workplace and they were used as a basis for content and knowledge provision within extended organizations.

The objective of K-NET project [28] is to support knowledge sharing and reuse within an organisation. Unlike the CKP tool, as well as the whole IntelLEO project, this project does not consider the specificities of learning in an extended organisation.

The European Integrating Project MATURE [29] is focused on providing technology that would allow an enterprise to make a significant shift in its organisational (learning) culture and move towards enterprise 2.0, which is characterized by enhanced collaboration and a culture...
of employee participation [20]. Unlike the CKP tool, this project does not provide means for achieving the goals of harmonization of personal and organizational goals, neither supports learning and knowledge building activities in an extended organizations. The latest trends in knowledge management are about using social software for conversations and collaboration, for knowledge elicitation, creation and sharing, for identifying experts and getting access to expert opinions worldwide. However, despite their numerous positive sides, social software tools also have one major drawback: the knowledge (i.e. knowledge objects) they capture is not accessible for machine processing. Therefore, there is a need for enriching these tools with formal semantics that can be leveraged by machines for supporting learning and knowledge building activities. Specifically, there is a need for annotating semantically knowledge objects created using social software tools.

There are a lot semantic annotation platforms, such as Action [21], AnnoTex [22], Self-teaching SVM struct [14], ASCUM [13] based on the domain ontologies. On the other hand, there are very popular bookmarking tools such as Delicious [18], Diigo [19] etc. The CKP tool implemented features for collaborative tagging in a form of an Internet browser plug-in (Firefox) are similar to other bookmarking systems at a glance. However, the key difference is that CKP can compute the relevance of the tagged resource with the learning goals and competences, and recommend learners to tag learning resources with them as well. This tool also differs from the other retrieval and content management services in a few ways. First, by using the CKP tool within the IntelLEO software solution, employees can store, annotate and (re-)discover heterogeneous resources (e.g., documents, discussions, blog posts, and wikis). The annotation is done automatically by using the concepts of appropriate domain-specific ontologies. Secondly, it allows employees to find job-specific experiences and “know-how” (in the form of e.g., annotated wiki pages, blog posts, discussions) that are not freely available on the Web. These can originate from a member of the extended organization, a colleague from the same organization, or can be a documented self experience. Annotations of these resources with the concepts from specific domain ontologies facilitate their discovery and retrieval. The IntelLEO CKP tool also aims at addressing well known drawbacks of traditional search paradigms related to difficulties in finding relevant information [12] by improving existing search interfaces with semantic search capabilities, thus allowing the search to be based on domain topics and not only keywords. Finally, the CKP tool offers a measuring the semantic similarity between learning resource and the specific learning goal which is a novel approach. It is a very important in determining the relevance of knowledge asset related to learning goal/competency that has to be achieved

VI. CONCLUSION AND FUTURE WORK

This paper presented the development process of the Content and Knowledge Provision tool by using the SOA concept. The tool is designed as a core service of IntelLEO, a new workplace learning paradigm being developed within an FP7 research project in the area of technology-enhanced learning. The concept of CKP tool requires technologies to support sharing, harmonization, building, and extension of content/knowledge among individuals, industries and universities, and effective combination of content and organizational knowledge systems (at both universities and workplaces in organizations). The management of content and knowledge has a key role for both collaborative LKB and harmonization of individual and organizational objectives.

This interface of the CKP tool includes all the advantages of the bookmarklets features. However, this service is much more than a simple bookmarking system or a mere Learning Management system; it combines all features in one single solution, thus improving exponentially the way those facilities can support workplace learning. The objective was to formally represent the semantics of the knowledge captured through social software tools, so that it can be leveraged by machines for supporting LKB activities in extended organizations.

In the future, we plan to implement the integration of the CKP tool with the Delicious bookmarking tool so that users can automatically save a bookmark to both Delicious and the CKP tool. Additionally, this integration should provide users with the ability to have recommended concepts and most frequently used tags, as well as the ability to search through colleagues’ bookmarks (both on Delicious and CKP).

ACKNOWLEDGMENT

This publication was partially supported/co-funded by the European Community/European Union/EU under the Information and Communication Technologies (ICT) theme of the 7th Framework Program for R&D (FP7). This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

REFERENCES


