Improving Knowledge Retrieval in Digital Libraries Applying Intelligent Techniques

Antonio Martín and Carlos León
Polytechnic School, Sevilla University, Avda. Reina Mercedes, Seville, Spain

Keywords: e-learning Systems, Ontology, Case-based Reasoning, Knowledge Management, Intelligent Agents.

Abstract: Nowadays an enormous quantity of heterogeneous and distributed information is stored in the digital University. Exploring online collections to find knowledge relevant to a user's interests is a challenging work. The artificial intelligence and Semantic Web provide a common framework that allows knowledge to be shared and reused in an efficient way. In this work we propose a comprehensive approach for discovering E-learning objects in large digital collections based on analysis of recorded semantic metadata in those objects and the application of expert system technologies. We have used Case Based-Reasoning methodology to develop a prototype for supporting efficient retrieval knowledge from online repositories. We suggest a conceptual architecture for a semantic search engine. OntoUS is a collaborative effort that proposes a new form of interaction between users and digital libraries, where the latter are adapted to users and their surroundings.

1 INTRODUCTION

Digital repositories enable users to interact effectively with information distributed across a network: publications, forms, guides, educational objects, legislation, etc. Access to these collections poses a serious challenge, however, because present search techniques based on manually annotated metadata and linear replay of material selected by the user do not scale effectively or efficiently to large collections. In the traditional search engines the information is treated as an ordinary database that manages the contents. The result generated by the current search engines is a list of Web addresses that contain or treat the pattern. The useful information buried under the useless information cannot be discovered. It is disconcerting for the end user and sometimes it takes a long time to search for needed information. Despite large investments and efforts have been made, there are still a lot of unsolved problems.

Thus, it is necessary to develop new intelligent and semantic models that offer more possibilities. Ontologies assist the extraction of concepts from unstructured textual documents and E-learning objects by serving as a source of knowledge in large digital libraries.

There is a lot of researches on applying these new technologies into current information retrieval systems, but no research addresses Artificial Intelligence (AI) and semantic issues from the whole life cycle and architecture point of view. The study (Jimeno-Yepes, 2010) presents a system, which uses an ontology query model to analyse the usefulness of ontologies in effectively performing document searches and proposes an algorithm to refine ontologies for information retrieval tasks with preliminary positive results. The study (Sasaki, 2005) presents a formulation and case studies of the conditions for patenting content-based retrieval processes in digital libraries, especially in image libraries. (Diaz-Galiano, 2009) uses a medical ontology to improve a Multimodal Information Retrieval System by expanding the user's query with medical terms. This study (Chen, 2008) combines swarm intelligence and Web Services to transform a conventional library system into an intelligent library system with high integrity, usability, correctness, and reliability software for readers. This research (Toledo, 2011) proposes organizational memory architecture, and annotation strategies based on domain ontologies to retrieve information through natural language queries.

Although search engines have developed effective searches, information overload obstructs precise searches. Our work differs from related

Martín A. and León C..
Improving Knowledge Retrieval in Digital Libraries Applying Intelligent Techniques.
DOI: 10.5220/0005159604450453
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2014), pages 445-453
ISBN: 978-989-758-049-9

projects in that we build ontology-based contextual profiles and we introduce an approach used metadata-based in ontology search and expert system technologies. This work presents an intelligent approach for optimize a search engine in the specific domain of university storehouses. It incorporates Semantic Web and AI technologies to enable not only precise location of public resources but also the automatic or semi-automatic learning.

The contribution has been divided into next sections. In the first section, short descriptions of important aspects in digital library domain and semantic interoperability, the research problems and current work in it are reported. The second section focuses on the Ontology design process and provides a general overview about our prototype architecture. Then we summarize its main components and describe how can interact AI and Semantic Web to enhancement a search engine. Finally we present the results of our on going work on the adaptation of the framework and we outline the future works.

2 DIGITAL LIBRARY ON E-LEARNING DOMAIN

Repositories and digital archives are privileged area for the application of innovative, knowledge intensive services that provide a flexible and efficient method for searching information and guarantee the user with a set of results actually related to his/her interest. These network information systems support search and display of items from organized collections. A Digital Library (DL) is an electronic storage and access environment for information and knowledge stored in the digital format either locally in a library, in a group of networked libraries, or at remote location. Reuse this knowledge is an important area in this domain. The Semantic Web provides a common framework that allows knowledge to be shared and reused across user's community (Sure & Studer, 2005).

Seville University institutional repository is dedicated to the production, maintenance, delivery, and preservation of a wide range of high-quality networked resources for scholars, and students at University and elsewhere. DL means an integrated set of services for capturing, cataloguing, storing, searching, protecting, and retrieving knowledge. It comprises digital collections, services, and infrastructure to support lifelong learning, research, scholarly communication, preservation, etc.

Our aim here is thus to contribute to a better knowledge retrieval in the institutional repositories field. This scheme is based on the next principles: knowledge items are abstracted to a characterization by metadada description which is used for further processing. This begets new challenges to docent community and motivates researchers to look for intelligent information retrieval approach and ontologies that search and/or filter information automatically based on some higher level of understanding are required. To reach these goals we need to consider information interoperability. In other words the capacity of different information systems, applications and services to communicate, share and interchange data, information and knowledge in an effective and precise way, as well as to integrate with other systems, applications and services in order to deliver new electronic products and services.

European initiatives, such as interoperability between public services, require establishing collaborative semantic repositories among public and private sector organizations. In this paper we study architecture of the search layer in this particular dominium, a web-based catalogue for the University of Seville. The hypothesis is that with a case-based reasoning expert system and by incorporating limited semantic knowledge, it is possible to improve the effectiveness of an information retrieval system (Sun and Finnie, 2004). More specifically, the objective is to explore and understand the requirements for rendering semantic search in an institutional repository and investigate from a search perspective possible intelligent infrastructures form constructing decentralized digital repositories where no global schema exists.

2.1 Interoperability Requirements

Particularly we require Semantic Interoperability, which is one of the key elements of the programme to support the set-up of the European E-Services In June 2002; European heads of state adopted the Europe Action Plan 2005 at the Seville summit. It calls on the European Commission to issue an agreed interoperability framework to support the delivery of European E-Government services to users and citizens. This recommends technical policies and specifications for joining up public administration information systems across the EU.

These aspects are the pillars to support the European delivery of E-Services of the recently adopted European Interoperability Framework (EIF) (SEC, 2003) and its Spanish equivalent (MAP, 2014). European Institutions and Agencies should use the European interoperability framework for their operations with each other and with users, enterprises and administrations in the respective EU Member States (EIF, 2014).

The ISO/IEC 2382 Information Technology Vocabulary defines the aspects of interoperability as a general concept or approach cover technical, semantic, and organisational issues, usually referenced as interoperability layers, Figure 1.

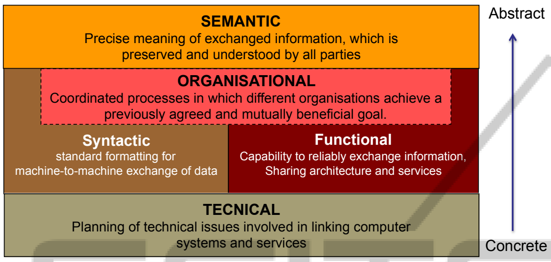

Figure 1: Abstraction layers interoperability.

Interoperability is conceived on different main abstraction levels:

- Organisational interoperability level: processes, defined as workflow sequences of tasks, integrated in a service-oriented environment.
- Technical interoperability level: signals, low-level services and data transfer protocols.
- Semantic interoperability level: information in various shared knowledge representation structures such as taxonomies, ontologies, or topic maps. Semantic interoperability shared vocabulary, and its associated links to an ontology, which provides the basis for machine interpretation and understanding of the logic of the message.

Exchanging normalized data is a prerequisite for semantic interoperability and refers to the packaging and transmission mechanisms for data. Two or more entities achieve interoperability when they are capable of communicating and exchanging data, which concerns to specified data formats and communication protocols. In this study we have focused our work in semantic interoperability analysis. This area implies the collaboration of many actors, such as local repositories. For this purpose we use ontologies and semantic approach, which enable reusing of existing domain knowledge and its further retrieval, providing a contextual framework enabling unambiguous communication of information represented.

However, semantic interoperability problems emerge as these organizations may differ in the terms and meanings they use to communicate, express their needs and describe resources they make available to each other. We must bear in mind that interoperability framework is, therefore, not a static document and may have to be adapted over time as technologies, standards and administrative requirements change. In the next sections we establish the base of all these aspects in our platform OntoUS.

3 ONTOUS ARCHITECTURE

In order to support semantic retrieval knowledge in university institutional repositories we develop a prototype named OntoUS based on ontologies and expert system technologies. OntoUS can be part of a bigger framework of interacting global information networks including e. g. other digital libraries, scientific repositories and commercial providers, and relies as much as possible on standards and existing building blocks as well as be based on web standards. The architecture of our system is shown in Figure 2, which mainly includes three parts: intelligent user interface, ontology knowledge base, and the search engine.

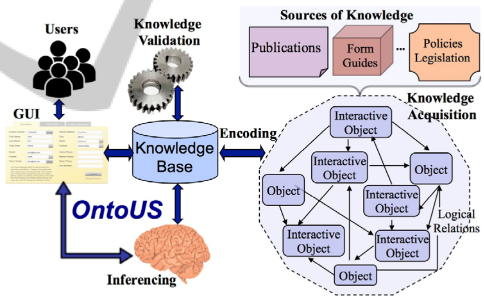

Figure 2: System architecture of OntoUS.

The proposed architecture is based on our approach to information retrieval in an efficient way by means of metadata characterizations and domain ontology inclusion. It implies to use ontology as vocabulary to define complex, multi-relational case structures to support the Case Based-Reasoning (CBR) processes. Our system works comparing objects that can be retrieved across heterogeneous repositories and capturing a semantic view of the world independent of data representation. OntoUS system uses its internal knowledge bases and inference mechanisms to process information about the electronic resources in Seville University repositories. Ontology knowledge base is the kernel part for semantic retrieval information. At this stage we consider to use ontology as vocabulary for defining the case structure like attribute-value pairs.

Ontology is the knowledge structure, which identify the concepts, property of concept, resources, and relationships among them to enable share and reuse of knowledge that are needed to acquire knowledge in the specific search domain.

Metadata elements have been effectively used for providing a richer representation of digital objects and collections. The metadata descriptions of the resources and repository objects (cases) are abstracted from the details of their physical representation and are stored in the Case Base.

Inference Engine contains a CBR component that automatically searches for similar queries-answer pairs based on the knowledge that the system extracted from the questions text. Case Base has a memory organization interface that assumes that whole case-base can be read into memory for the CBR to work with it. We used a CBR shell, software that can be used to develop several applications that require cased-based reasoning methodology. In this work we analysed the CBR object-oriented framework development environments JColibri (GAIA, 2009). This framework work as open software development environment and facilitate the reuse of their design as well as implementations. The CBR engine uses an evaluation function to calculate the new case ranking, and the answered question updates the query and the rankings in the displays. The questions are ranked according to their potential for retrieval and matching.

Advanced conversational user interface interacts with users to solve a query, defined as the set of questions selected and answered by the user during conversation. Interface is designed and developed to improve communication between humans and the platform. In our system the user interacts with the system to fill in the gaps to retrieve the right cases. Also we have implemented a context interface, which allows retrieving cases enough to satisfy a SQL query.

3.1 Case-based Reasoning

CBR is widely discussed in the literature as a technology for building information systems to support knowledge management, where metadata descriptions for characterizing knowledge items are used. We have chosen the framework jColibri a java-based configuration that supports the development of knowledge intensive CBR applications and help in the integration of ontology in them. In our CBR application, problems are described by metadata concerning desired characteristics of an institutional resource, and the solution to the problem is a pointer to a resource described by metadata. These characterizations are called cases and are stored in a Case Base. CBR case data could be considered as a portion of the knowledge (metadata) about an OntoUS object. Every case contains both index with the association terms of the ontology and the relation documents residing on the repository network.

The development of a quite simple Case-Based Reasoning application already involves a number of steps, such as collecting case and background knowledge, modelling a suitable case representation, defining an accurate similarity measure, implementing retrieval functionality, and implementing user interfaces. Compared with other AI approaches, CBR allows to reduce the effort required for knowledge acquisition and representation significantly, which is certainly one of the major reasons for the commercial success of CBR applications.

Case-Based Reasoning is a problem solving paradigm that solves a new problem, in our case a new search, by remembering a previous similar situation and by reusing information and knowledge of that situation. Reasoning cycle may be described by the following steps processes, Figure 3:

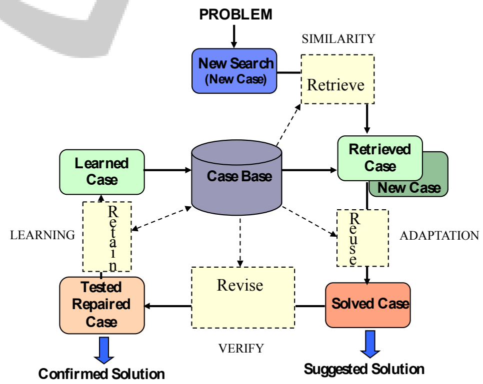

Figure 3: Case Based Reasoning Cycle in OntoUS.

- Retrieval. The system retrieves the closest-matching cases stored in a case base. Main focus of methods in this category is to find similarity between cases.
- Reuse: a complete design where case-based and slot-based adaptation can be hooked is provided. If appropriate, the validated solution is added to the case for use in future problem solving.
- Revise the proposed solution if necessary. Since the proposed result could be inadequate, this process can correct the first proposed solution. It should be noted that the differences in adaptation power

depend on how well the domain is understood.

- Retain the new solution as a part of a new case. The solution is validated through feedback from the user or the environment. This process enables CBR to learn and create a new solution that should be added to the knowledge base.

The Open Source JColibri1 system provides a framework for building CBR systems based on state-of-the-art Software Engineering techniques. Our motivation for choosing jColibri framework is based on a comparative analysis between it and other frameworks, designed to facilitate the development of CBR applications. jColibri enhances the other CBR shells: CATCBR, CBR*Tools, IUCBRF, Orenge, in several aspects: open source framework and Java implementation, which is one of our main requirements with respect to the easy integration in the OntoUS system implemented in J2EE environment.

Another decision criterion for our choice is connected with the fact that jColibri affords the opportunity to incorporate ontology in the CBR application to use it for case representation and content-based reasoning methods to assess the similarity between them. By providing easy to use model generation, data import, similarity modelling, explanation, and testing functionality together with comfortable graphical user interfaces, the tool enables even CBR novices to rapidly create their first CBR applications. Nevertheless, at the same time it ensures enough flexibility to enable expert users to implement advanced CBR applications.

3.2 Ontology Development

The main objective of our system is to improve the modelling of a semantic coherence for allowing the interoperability of different modules of environments dedicated to digital university. We have proposed to use ontology together with CBR in the acquisition of an expert knowledge in the specific domain. The primary information managed in the OntoUS domain is metadata about institutional resources, such as guides, publications, forms, digital services, etc. We need a vocabulary of concepts, resources and services for our information system described in the scenario requires definitions about the relationships between objects of discourse and their attributes (Taniar and Wenny Rahayu, 2006). OntoUS project contains a collection of codes, visualization tools, computing resources, and data sets distributed across the grids, for which we have developed a well-defined ontology using RDF language. RDF is used to define the structure of the metadata describing resources. Our ontology can be regarded as triplet OntoSearch:={profile, collection, source, relation) where profiles represent the user kinds, collection contains all the services and resources of the institutional repository and matter cover the different information sources: electronic services, official web pages, publications, guides, etc., and relation is a set of relationships intended primarily for standardization across ontologies. A detailed picture of our effort in designing this ontology is available in the Figure 4.

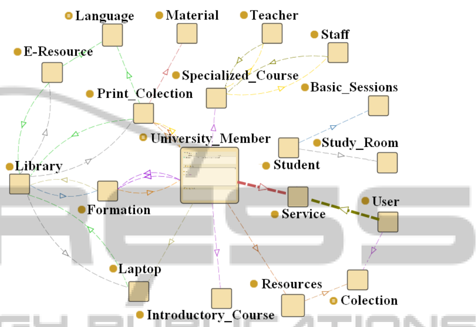

Figure 4: Ontological distributed environment.

This shows the high level classification of classes to group together OntoUS resources as well as things that are related with these resources. Profile ontology includes several attributes like Electronic_Resources, Digital_Collections, Publication_Catalogue, Public_Services, etc. We choose Protégé as our ontology editor, which supports knowledge acquisition and knowledge base development (Protégé, 2013). It is a powerful development and knowledge-modelling tool with an open architecture. Protégé uses OWL and RDF as ontology language to establish semantic relations. For the construction of the ontology of our system, firstly we determine the domain and scope of the ontology: Publications Catalogue, Web Sites, Electronic Resources, etc. Also ontology must be adapted to needs of user kinds. Second we enumerate important terms in the ontology. It is useful to write down a list of all terms we would like either to make statements about or to explain to a user. Then we define the classes and the class hierarchy. The ontology and its sub-classes are established according to the taxonomies profile.

In order to relate ontology classes to each other, we defined our own meaningful properties for the ontology and we defined a class hierarchy associated with meaningful properties. Slots can have different facets describing the value type, allowed values, the number of the values (cardinality), and other features of the values the slot can take. In the following we give a short RDF description that defined the concept of the user teacher that is a subclass of Member_Community_University.

```
<rdf:Description rdf:about="#Teacher">
<rdfs:comment rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
Teacher profile for affiliated colleges
   </rdfs:comment>
     <rdfs:subClassOf rdf:resource=
   "#Members of the University community"/>
   <rdf:type rdf:resource=
   "http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
```

The last step is to provide a conversational CBR system to retrieve the requested metadata satisfying a user query we need to add enough initial instances and item instances to knowledge base. 10.000 cases were collected for user profiles and their different resources and services. This is sufficient for our proof-of-concept demonstration, but would not be sufficiently efficient to access large resource sets. Each case contains a set of attributes concerning both metadata and knowledge. However, our prototype is currently being extended to enable efficient retrieval directly from a database, which will enable its use for large-scale sets of resources. As a plus, domain specific rules defined by domain experts (manually or by tools) can infer more complex high-level semantic descriptions, for example, by combining low-level features in local repositories. On one hand, the rules can be used to facilitate the task of resource annotation by deriving additional metadata from existing ones.

We come to a process for addressing complex relations between ontologies. As mentioned in previous sections, relations among ontologies can be composed as a form of declarative rules, which can be further, handled in inference engines. In our approach, we choose to use the Semantic Web Rule Language (SWRL), which is based on a combination of OWL DL and OWL Lite with the Case-Based reasoning sublanguages, to compose declarative search rules.

4 VISUAL INTERFACE TO EARLIER RETRIEVAL

The understanding about digital libraries and repositories is quite different according to its specific users. OntoUS monitors user's tasks, anticipates search-based information needs, and proactively provide users with relevant information. The objective of profile intelligence has focused on creating of user profiles: Staff, Administrator, and Visitor. The user interface helps to user to build a particular profile that contains his interest search areas in the university repositories domain. In an intelligence profile setting, people are surrounded by intelligent interfaces merged, thus creating a computing-capable environment with intelligent communication and processing available to the user by means of a simple, natural, and effortless human-system interaction. The user enters query commands and the system asks questions during the inference process. Besides, the user will be able to solve new searches for which he has not been instructed, because the user profiles what he has learnt during the previous searchers.

This configuration contains the user requirements most typically described the relative needs, tasks, and goals of the user for an individual search. For this a statistical analysis has been done to determine the importance values and establishing specified user requirements. A schematic of the architecture is show in Figure 5.

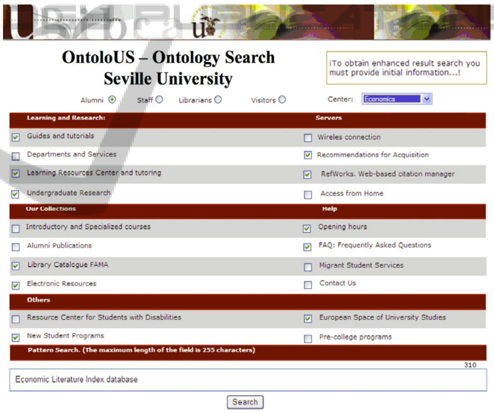

Figure 5: User Profiles, Graphical User interface.

The advantage of CBR is that users need only input text partially describing the search and then the system can assist in further complete the problem description in an interactive conversation style. The following guidelines for CBR design were proposed: reuse questions, order context questions before detail questions, eliminate questions that do not distinguish cases, ask for only one thing in a question, and use a similar, short number of questions per case. The user begins the search devising the starting query. After searching, some resources are returned as results. The results include a list of web pages with titles, a link to the page, and a short description showing where the keywords have matched content within the page.

5 RETRIEVAL PROCESS OF SIMILAR CASES

CBR systems typically apply retrieval and matching algorithms to a case base of past problem-solution pairs. CBR is based on the intuition that new searches are often similar to previously encountered searches, and therefore, that past results may be reused directly or through adaptation in the current situation. Retrieval processes get back information from the case library a set of potentially useful cases, all of which partially match the new situation. Retrieval process identifies the features of the case with the most similar query. Our Inference Engine contains the CBR component that automatically searches for similar queries-answer pairs based on the knowledge that the system extracted from the questions text. The system uses similarity metrics to find the best matching case. Similarity retrieval expands the original query conditions, and generates extended query conditions, which can be directly used in knowledge, Figure 6.

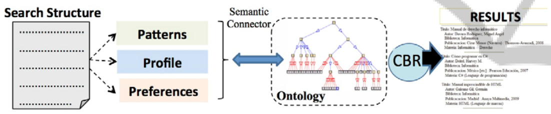

Figure 6: Retrieval cases process.

The use of structured representations of cases requires approaches for similarity assessment that allow to compares two differently structured objects, in particular, objects belonging to different object classes. Retrieval strategy used in our system is euclidean approach (Finnie and Zhaohao, 2002). This approach involves the assessment of similarity between stored cases and the new input case, based on matching a weighted sum of features. Euclidean distance is basis of many measures of similarity and dissimilarity, which is usually the right measure for comparing cases. The distance between ranking vectors $case_1$ and $case_2$ is defined as follows:

$$similarity(Case_i, Case_R) = \sqrt{\sum_{i=1}^{n}(Case_i - Case_i)^2} \qquad (1)$$

Euclidean distance is the square root of the sum of squared differences between corresponding elements of the two vectors. In our system euclidean distance is used to compare search results across variables. Each row of the matrix is a vector of n numbers, where n is the number of variables. We evaluate the distance, or in this case the similarity between any pair of rows. An important advantage of similarity-cased retrieval is that if there is no case that exactly matches the user's requirements, this can show the cases that are most similar to her query.

6 EVALUATION AND PROOFS

Experiments have been carried out in order to evaluate the effectiveness of run-time ontology mapping. The main goal has been to check if the mechanism of query formulation, assisted by an agent, gives a suitable tool for augmenting the number of significant documents, extracted from the Seville University institutional repository, to be stored in the CBR. For our experiments we considered 150 users with different profiles. So that we could establish a context for the users, they were asked to at least start their essay before issuing any queries to system. They were also asked to look through all the results returned by OntoUS before clicking on any result. In each experiment we report the average rank of the user-clicked result for our baseline system, Google and for our search engine OntoUS. Then we calculated the rank for each retrieval document by combining the various values and comparing the total number of extracted documents and documents consulted by the user, Figure 7.

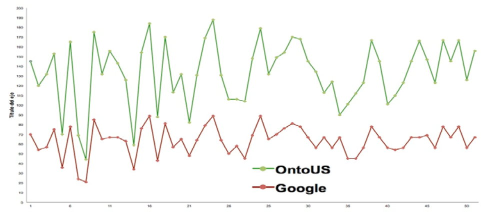

Figure 7: Comparative of valid pages percentage.

In our study domain we can observe the best final ranking was obtained for our prototype OntoUS and an interesting improvement over the performance of Google. Our system performs satisfactorily with about a 94.6% rate of success in real cases. Another important aspect of the design and implementation of an intelligent system is determination of the degree of speed in the answer that the system provides. During the experimentation, heuristics and measures that are commonly adopted in information retrieval have been used, Figure 8.

While the users were performing these searches, an application was continually running in the background on the server, and capturing the content of queries typed and the results of the searches.

Statistical analysis has been done to determine the importance values in the results. Previous figure shows a sample plot of these parameters that was collected as a part of the experiment. We can establish that OntoUS speed in our DL domain improves the proceeding time and the average of the traditional search engine.

Figure 8: OntoUS search analysis report.

The results for OntoUS are 21.5% better than proceeding time and 14.9% better than executing time searches/sec in the traditional search engines.

7 CONCLUSIONS

Our study addresses the main aspects of a Semantic Web information retrieval system architecture trying to answer the requirements of the next-generation Semantic Web user. For this purpose we presented a system based on ontology and AI architecture for knowledge management in the Seville repositories. This scheme is based on the next principle: knowledge items are abstracted to a characterization by metadata description witch is used for further processing.

We have been working on the design of entirely ontology-based structure of the case and the development of our own reasoning methods in jColibri to operate with it. It introduced a prototype web-based CBR retrieval system OntoUS, which operates on an RDF file store. Furthermore an intelligent agent was illustrated for assisting the user by suggesting improved ways to query the system on the ground of the resources in Seville University Repositories according to his own preferences, which come to represent his interests. Finally the study analyses the implementation results, and evaluates the viability of our approaches in enabling search in intelligent-based digital repositories.

Future work will concern the exploitation of information coming from others institutional repositories and digital services and further refine the suggested queries, to extend the system to provide another type of support, as well as to refine and evaluate the system through user testing. It is also necessary the development of an authoring tool for user authentication, efficient ontology parsing and real-life applications

REFERENCES

Chen, L., (2008). Design and implementation of intelligent library system. *Library Collections, Acquisitions, and Technical Services, Vol. 32, Issues 3–4, 2008, 127-141.*

Diaz-Galiano, M. C., Martin-Valdivia, M. T. & Urena, L.A. (2009), Query expansion with a medical ontology to improve a multimodal information retrieval system. *Computers in Biology and Medicine, Vol. 39, Issue 4.*

European Interoperability Framework (EIF) Version 2. (2004) [cited 21-01-2014]; Available from: ec.europa.eu/isa/strategy/doc/annex_ii_eif_en.pdf.

Finnie G and Sun, Z. (2002). Similarity and metrics in case-based reasoning. *Int J Intelligent Systems 17 (3), 273-287.*

GAIA - Group for Artificial Intelligence Applications, (2014). *jCOLIBRI project*, http://gaia.fdi.ucm.es/grupo/projects/. Complutense University of Madrid.

Jimeno-Yepes, A., Berlanga-Llavori, R. & Rebholz-Schuhmann, D. (2010). Ontology refinement for improved information retrieval, Information Processing & Managemen. *Volume 46, Issue 4, Semantic Annotations in Information Retrieval.*

Ministerio de Administraciones Públicas (2014). *Aplicaciones utilizadas para el ejercicio de potestades. Criterios de Seguridad, Normalización y Conservación:* 22/01/2014 Version. <http://www.csi.map.es/csi/criterios/index.html>.

PROTÉGÉ, (2013). *The Protégé Ontology Editor and Knowledge Acquisition System.* <http://protege.stanford.edu/>.

Sasaki, H., & Kiyoki, Y. (2005). A formulation for patenting content-based retrieval processes in digital libraries, *Information Processing & Management, Volume 41, Issue 1, Pages 57-74.*

SEC (2003) 801. *Commission Staff Workking Paper: linking up Europe, the importance of interoperability for egovernment services* europa.eu.int/ISPO/ida/export/files/en/1523.pdf.

Sun, Z. and Finnie G. (2004): *Intelligent Techniques in E-Commerce: A Case-based Reasoning Perspective.* Heidelberg: Springer-Verlag.

Sure, Y., and Studer, R., (2005). Semantic Web technologies for digital libraries. *Library Management Journal, Emerald, vol. 26.*

Taniar, D., Wenny Rahayu, J. (2006). *Web semantics and ontology*. Hershey, PA: Idea Group Pub, 2006.

Toledo, C. M. & Ale M. A. (2011). An Ontology-driven Document Retrieval Strategy for Organizational Knowledge Management Systems, *Electronic Notes in Theoretical Computer Science, Vol. 281, 21-34.*