Interaction Design and Architecture(s) Journal - IxD&A, N.23, 2014, pp. 38-49

Cloud Computing and Multi Agent System to improve Learning Object Paradigm

Fernando De la Prieta, Ana B. Gil, Sara Rodríguez-González, Juan M. Corchado

University of Salamanca, Department of Computer Science and Automation Control, Plaza de la Merced s/n, 37007, Salamanca, Spain
{fer, abg, srg, corchado}@usal.es

Abstract. The paradigm of Learning Object provides Educators and Learners with the ability to access an extensive number of learning resources. To do so, this paradigm provides different technologies and tools, such as federated search platforms and storage repositories, in order to obtain information ubiquitously and on demand. However, the vast amount and variety of educational content, which is distributed among several repositories, and the existence of various and incompatible standards, technologies and interoperability layers among repositories, constitutes a real problem for the expansion of this paradigm. This study presents an agent-based architecture that uses the advantages provided by Cloud Computing platforms to deal with the open issues on the Learning Object paradigm.

Keywords: Learning Objects, digital repositories, information retrieval, cloud computing

1 Introduction

The Learning Object (LO) paradigm is one of the main advances within the field of reutilization of educational resources. This new paradigm initiates a new set of technological goals related to the new life-cycle of creating educational experiences. An LO can be described as a digital, self-contained and reusable entity with a clear educational purpose, composed of at least three internal components: content, learning activities, and contextualization elements. In short, it refers to practically any educational resource (lesson, task, graph, subject, etc.) that can be described by means for metadata.

Formally, an LO is defined [15] as *any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning*. There is a clear consensus that an LO must be the minimal reusable self-contained unit of learning content with a specific objective [3][18]. The paradigm is based in the fact that any education resource can be described by means of metadata, independently of its topic, type, format, size, etc. The encapsulation of education resources in the form of metadata makes their digital distribution and subsequentr reutilization, possible because this metadata allows making a first approach to the educational resource. The metadata schema is standardized. In fact, there are currently many standards

[8][15][22], etc. The existence of standards facilitates the management of the resources, enabling the interoperability among systems that use compatible standards.

LOs are stored in specific digital libraries, called Learning Object Repositories (LOR). Currently there is a significant growth of LOR as part of the hidden web in large databases. These systems typically provide a web interface to allow a search of education resources through the metadata. One of the main characteristics of these LORs is their heterogeneity [11], which indicates that the interoperability among LORs is limited. However, to deal with this issue, they typically have a layer (interface) to make external access, and hence, interoperability, possible. External search agents (a client or another LOR) can access through web services. There are different standards or specifications that focus on this interoperability layer [19] [9] [16].

However, despite the theoretical advances made within this paradigm, reality shows that its implantation in real life is still limited [14]. From our point of view, there are two main problems:

- **Usability:** the data that the authors assign to each descriptor of the metadata (independent of the specific standard) is very important because this data is used for searches and if it is not correct, the results of the searches will be incoherent. Consequently, it is necessary to follow a traceable process from the creation of an educational resource to the creation of its metadata in order to establish a metadata structure that is consistent, relevant and interpretable. However, the existence of many standards, the interoperability among them, the difficulty in using the authoring tools and style of explanations by the authors, all exacerbate the problem.
- The **technology**, the heterogeneity of the repositories and their malfunction (as shown in the following section) constitute one of the main weakness of the paradigm.

This paper/article/study presents the evolution of the CLOR platform [5][6], which integrates the agent-based federated searcher AIREH [12] to produce a clear advantage in the context of the LO paradigm. To begin, CLOR (*Cloud-based Learning Object Repository*) is the current manifestation of a new generation of LOR because it is deployed into a cloud platform and takes advantage of this computational paradigm (non-SQL databases, unlimited storage, etc.). Additionally, AIREH (*Architecture for Intelligent Retrieval of Educational content in Heterogeneous Environments*) is a platform that enables federated searches among many LORs and integrates a recommender system [17].

This paper is structured as follows: the next section includes a study of the state of the art of current LOR. Starting from this study, section 3 presents an empirical study of the existing LORs. To solve the open issues, section 4 presents the proposed agent-based architecture. Finally, the preliminary results are included in section 5 and the conclusions in last section.

2 Open Issues – Learning Object Repositories

LOs are commonly stored in repositories, which are characterized by their heterogeneity. The deployment infrastructure is essentially either distributed or centralized

Taking into account that an LO is formed by a digital resource and its metadata, Frango *et al*. [11] identify four kinds of possible infrastructures as shown in **Fig. 1**: (i) centralized resources and centralized metadata, (ii) centralized resources and distributed metadata, (iii) distributed resources and centralized metadata and (iv) distributed resources and distributed metadata.

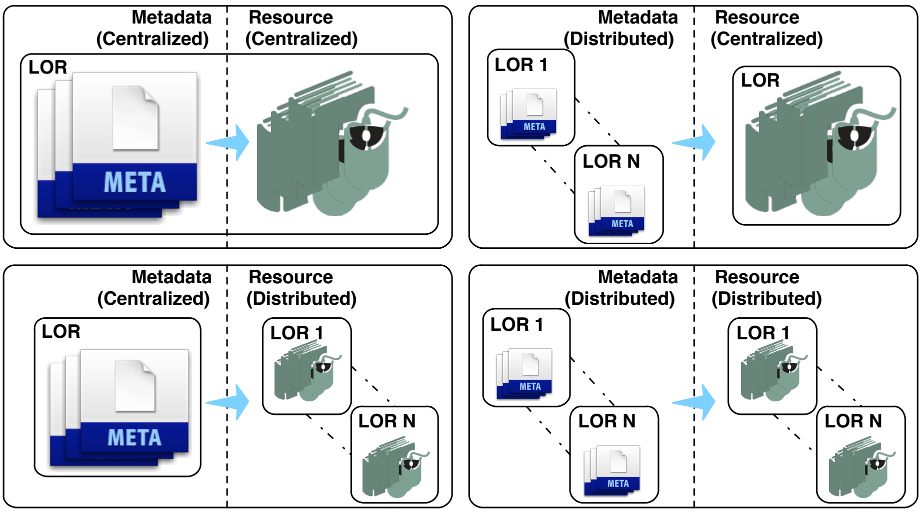

Fig. 1. Strategies to store LO in repositories.

Furthermore, Frango *et al*. distinguishe three kinds of storage strategies [11]: (i) *File-based,* which uses files with predefined formats and an index-based management; (ii) *Database-based,* which uses any kind of database, and is the most extended method; and (iii) *Persistent objects-based,* where the LO are stored as serialized objects.

With regard to metadata schemas, there are currently many standards such as DublinCore [8], IEEE LOM [15] and SCORM (Sharable Content Object Reference Model) [22]. The existence of standards facilitates the management of the resources, enabling the interoperability among systems that use compatible standards. However, reality shows that in some cases they are the problem, as many existing standards are not compatible among themselves. The ADLNet[1] initiative was developed in order to solve these problems and to coordinate the effort of metadata standards and, in general, the use of IT in the educational context. It is important to note the existence of metadata standards is not necessary in order to reuse contents.

[1] Advanced Distributed Learning Network. http://www.adlnet.org/

As previously explained, the last few years have seen a lot of research focusing on LOR. The research has resulted in the development of a number of standards, tools, methodologies and so on; in other words, heterogeneity. By isolating this internal heterogeneity (storage techniques and metadata schema), there are now interoperability layers which serve as a middleware layer between the repository and the clients (**Fig. 2**):

- **OAI-MPH** (***The Open Archives Initiative Protocol for Metadata harvesting***) **[19].** This is an open protocol that allows retrieving learning resources, thus ensuring interoperability. One of the main advantages is the design as technology-independent framework. This framework is extensively used by libraries, museums, etc. to exchange information about resources. However it is not specially design to exchange learning resources.
- **IMS DRI** (***IMS Digital Repository Interoperability***) **[16].** This is an abstract framework designed for the IMS Global Learning Consortium. It is based on the use of existing communication technologies and protocols to provide interoperability. However, this protocol is still in its incipient stages of development.
- **SQI** (***Simple Query Interface***) **[9].** This is formed by a set of abstract methods based on web services. These methods are not associated with any underlying technology. It is also neutral in terms of the format of results as well as query language. This web interface supports synchronous/asynchronous and stateful/stateless queries; authentication is based on a session with the aim of isolating the harvesting of contents from the management tasks.

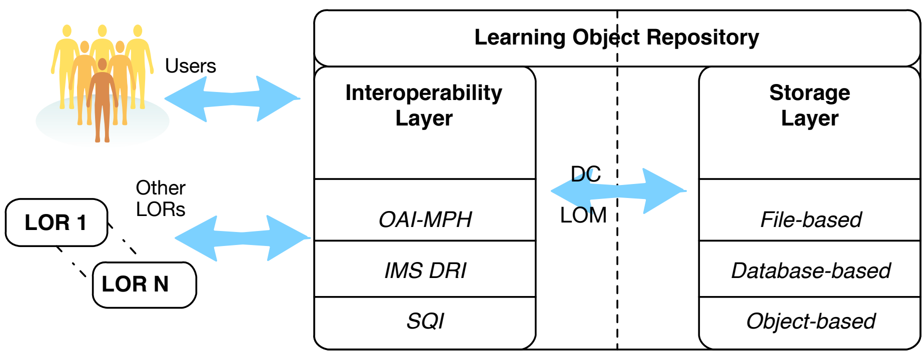

Fig. 2. Internal features of LOR

Although these interoperability layers may appear to reduce the main problems of the LO paradigm, related with the heterogeneity, there are still LORS that do not implement any abstraction layer that can encapsulate the internal logic of the repository. Consequently, the search process and LO harvesting is a slow process, which requires the manual intervention of users who must reuse the learning resources. Also, it is important to note that not only is the existence of metadata standards necessary in order to reuse contents, but the data that the authors assign to each descriptor is very important as well. It is therefore necessary to follow a traceable process from the creation of an educational resource to the creation of its metadata in order to establish a metadata structure that is consistent, relevant and interpretable. [2].

3 Empirical study of the LOR

The state of the art shows high heterogeneity in existing standards. However, a study of LOR was performed in order to analyze the real situation. In general, the systems in which this layer is included are prone to various problems such as:

- problems associated with the monolithic structure of LOR, which does not allow external management with the flexibility and power necessary to ensure easy interoperability, and dispersed and heterogeneous sources.
- the absence of automatic mechanisms that control the technical quality, semantics and syntax of LO, ensuring the correct specification of such LOs in any of the metadata schemas that describe them.

The study includes an analysis of the following LORs: *Acknowledge*, *Agrega*, *Ariadne*, *AriadneNext*, *CGIAR*, *EducaNext*, *LACLO-FLOR*, *LORNET*, *MACE*, *Merlot*, *Nime*, *OER Commons* and *Edna Online*. In the study 60 queries were performed on each LOR through an SQI layer that the repositories provide. All of them use IEEE LOM [15] as metadata schema and VSQL [2323] as query language. Additionally, the majority of them are stateless (65%), and all of them have synchronous interfaces, but only 4 have an asynchronous interface.

The test showed that 6 of the 14 repositories do not work or are unavailable, resulting in their removal from the scope of this study (Ariadne, AriadneNext, EducaNext, Nime and EdNa Online). MACE and LOCLO-FLOR produced an error in the authentication. After this step, this test was reduced to only four repositories Acknowledge, Agrega, LORNET and Merlot. The latter three are perfectly valid and all SQI methods work perfectly; however the repository Acknowledge only implements the essential methods to perform queries (note that SQI specification does not force the implementation of all methods of the specification).

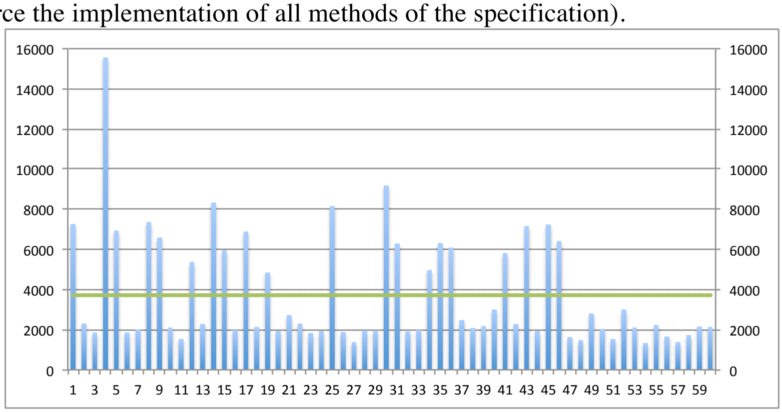

Fig. 3. Time average of 60 queries of a federated search in 4 repositories

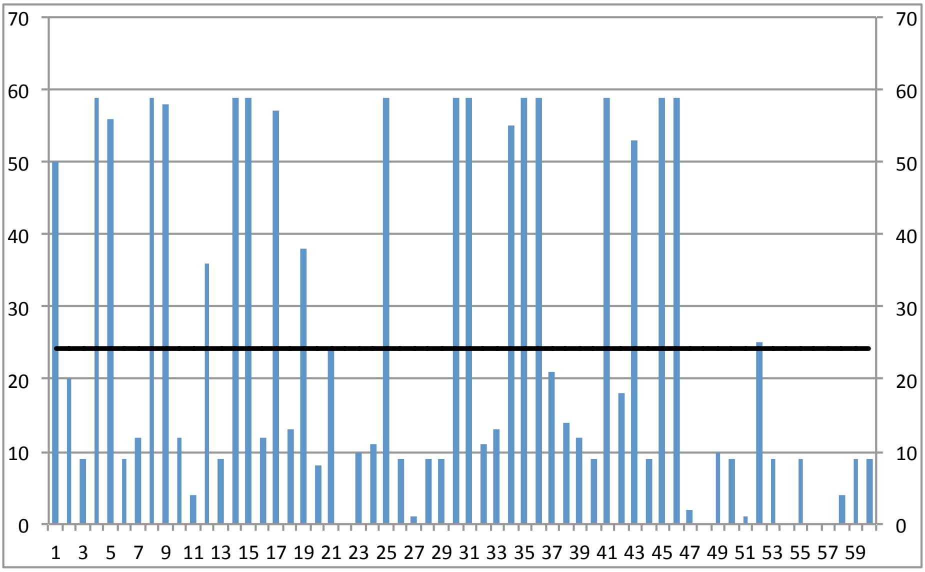

Fig. 4. Number of 60 queries results of a federated search in 4 repositories

As it is possible to observe, the performance of the LOR is not appropriate. In order to deal with this problem, new LOR architectures have to be proposed and developed. This new generation of LOR must ensure the availability of resources and interoperability, permitting federated searches from external clients.

3.1 An oportunity of Cloud Computing paradigm

Among the services in the context of the Internet, Cloud Computing has been recently emerging as a key paradigm of the present century. According to NIST[2] [21], *Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models*. This definition includes three levels of computational services (Software, Platform and Infrastructure).

This rapid growth is to a large extent tied to the more sophisticated developments that have been reached by related technologies. New possibilities at a technological level lead to the birth of a new concept, elasticity [4]. Cloud services are able to offer the same level of quality independently of instant demand. In practice, end users make use of Cloud services that are always available and unlimited. However, the services produced within the framework of CC only receive the amount of resources they need

[2] NIST, National Institute of Standards and Technology (http://www.nist.gov/)

to maintain a uniform level of quality while immediately responding to demand [25] [24].

This technological framework provides advanced tools and technologies that not only improve the storage of resources in terms of availaibility, time-response, storage space, etc., but also search and harvest learning resources over the cloud by means of use new services and their availability in order to provide a interoperability framework within this context. Among these technologies, we highlight the followings: Non-SQL databases and Storage Are Network.

4 AIREH Arquitecture

These problems require solutions that are adapted to heterogeneity. The solution should enable a centralized global search and the effective reuse of resources by the end user. As suggested in the introduction of the present study, AIREH and CLOR are integrated in order to establish a system that not only allows the federated search among several LOR, but can also retrieve and safely store the retrieved LO.

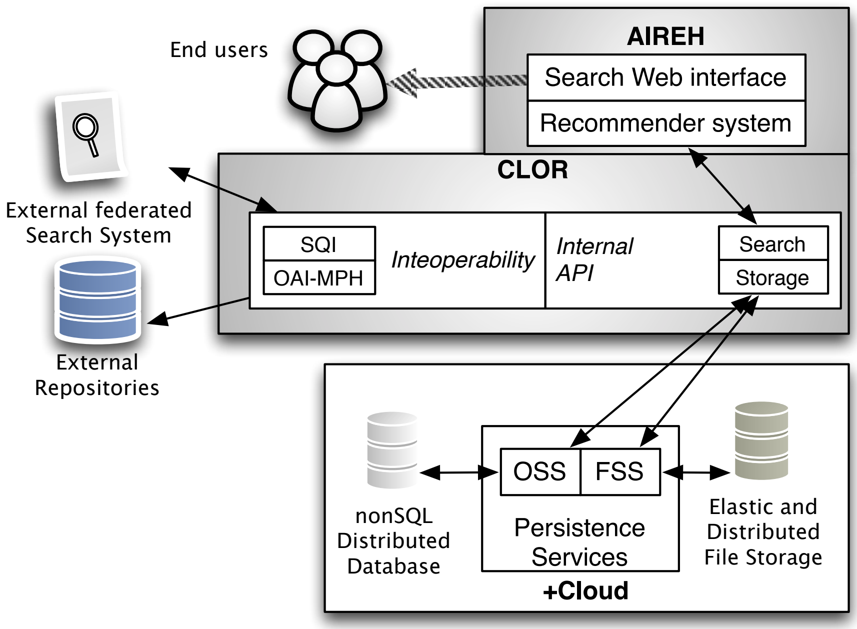

Fig. 5. CLOR and AIREH integration

The system components, which are shown in **Fig. 5**, are described as follows:

— AIREH (**Fig. 6**) unifies the search and retrieval of objects, thus facilitating the learning search process by filtering and properly classifying learning objects retrieved according to certain rules. The framework is based on a virtual organization [13] of intelligent agents that allows dealing with the heterogeneity of the environment. This agent-based architecture designed with GORMAS [6] will solve the problems of distribution, the integration of different repositories, the abstraction of the internal logic of each repository, and the classification, storage and retrieval of LOs, in a completely transparent way. In addition to adding capacities, such as simple scalability, and possible situations involving the use of new protocols, the architecture also adds internal logical repositories, as well as cataloging or heterogeneous applications designed to cover services related features. AIREH implements Case Based Reasoning (CBR)[10] [1] which uses previous search information to rank the items that best suit the needs of the application user based on previously obtained information. It uses the profile information of each user as well as their educational information (content-based filtering).

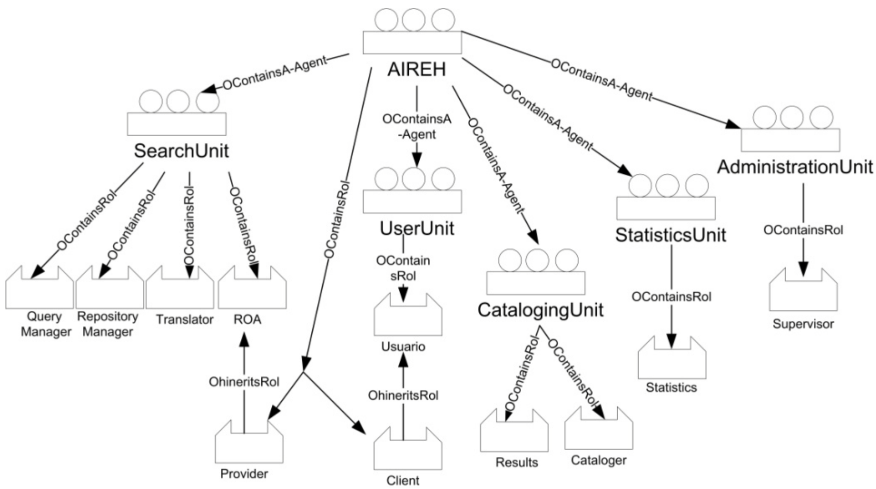

Fig. 6. AIREH functional structure diagram

— CLOR provides AIREH with the capacity to store the profile of each user as well as the persistence of retrieved LO (not only the metadata, but also the education resource). It is complemented with different interoperability layers, such as SQI or OAI-MPH, which will ensure the communication with other LORs and federated searches from external clients. It is framed at the platform level within Cloud services. Its main task is to encapsulate the communication with the lower layers of the Cloud platform that provide, firstly, the needed computational resources for storing the educational resources in the web service file storage system and, secondly, the metadata (in JSON format) associated with each resource into a non-SQL database. The main advantage is that it permits storing any kind of metadata independent of its structure or schema. Furthermore, queries about the LO will be performed very quickly thanks to the use of a document-oriented database [20].

- +Cloud platform [7] (**Fig. 7**) provides cloud capabilities, such as storage and databases. This platform is also designed using a virtual organization of intelligent agents, and is based on the Cloud Computing paradigm. This platform allows offering services at the PaaS and SaaS levels. The SaaS layer is composed of the management applications for the environment (virtual desktop, control of users, installed applications, etc.), and other more general third party applications that use the services from the PaaS layer. The components of this layer are the identity Manager, a File Storage System base on Web services, and an Object Storage Service, which provides a simple and flexible schemaless data base-oriented service.

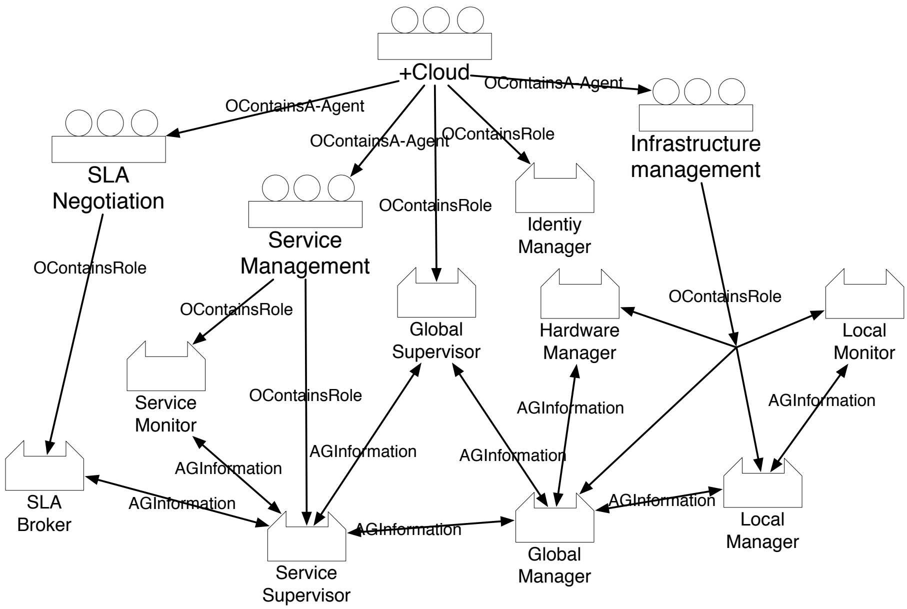

Fig. 7. +Cloud functional structure diagram

The main advantage of CLOR is that it allows both paradigms (LO and Cloud Computing) to coexist in a proactivity environment, where AIREH is able to retrieve information and execute the CBR cycle at high speed because it is situated in a high performance computational environment. At the sime time, +Cloud provides a specific environment persistence: non-relational databases to store metadata, and distributed storage to store resources.

5 Preliminary results

The proposed system was evaluated by performing a battery of tests to validate their efficiency in real environments. Evaluation metrics from information retrieval fields were adopted. The two most commonly used evaluation measures are precision

(Formula 1), which is the fraction of documents retrieved by the system that are also relevant to the query, and recall (Formula 2) (the fraction of the relevant documents present in the database that are retrieved by the system).

$$P_J = \frac{\sum_1^n R_i O_i}{n} \tag{1}$$

$$E_J = \frac{\sum_1^n R_i O_i}{\sum_{J=1}^m \sum_{i=1}^n R_{ij} O} \tag{2}$$

Fig. 8 shows the number of relevant LO retrieves per query compared with other repositories, and **Fig. 9** shows an average of time spent in retrieving LO per query.

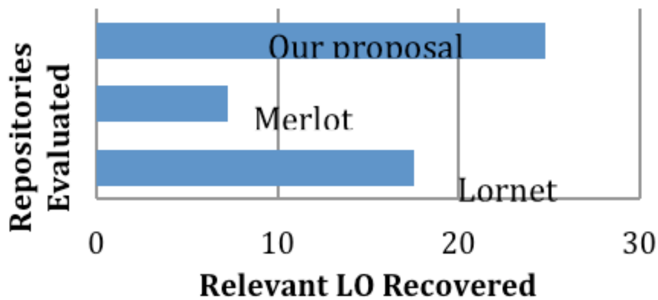

Fig. 8. Relevant LO recovered per query

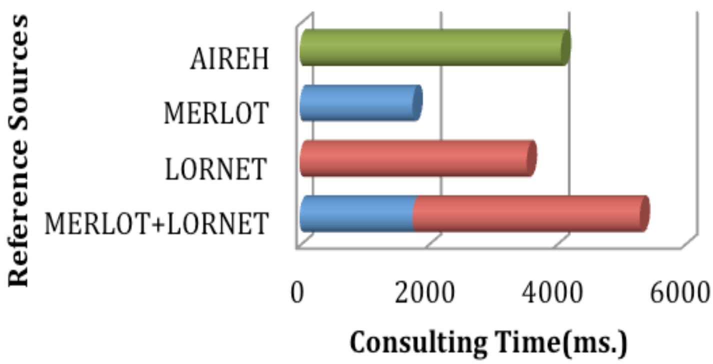

Fig. 9. Average time of queries

6 Conclusions

The first problem of the LO paradigm is the incoherence in the medatada. This incoherence is due to the fact that the labeling process, which is basically done by hand, generates documents with serious shortcomings, including many deficiencies related to the lack of key attributes in the description. This makes it difficult, or impossible in some cases, to study this aspect.

The second problem is the heterogeneity of the repositories and their malfunction. The proposed system tries to minimize this second problem, because it deals with existing open issues:

1. The proposed model allows dealing with the heterogeneity of current and future standards since it is based on a non-relational database.
2. Cloud computing paradigm makes it possible to offer services with the same level of quality, independent of demand.
3. The low linkage among components permits implementing many interoperability layers without needing to upgrade to other modules.

The integration of AIREH allows the results to be filtered according to certain parameters related to the quality of the retrieved metadata documents; the user context information is then integrated into the use of the LOs.

Acknowledgments. This work is supported by the Spanish government (MICINN) and European FEDER funds, project iHAS: Intelligent Social Computing for Human-Agent Societies (TIN2012-36586-C03-03)

References

1. Bajo J., Corchado, JM. Evaluation and monitoring of the air-sea interaction using a CBR-Agents approach. Case-Based Reasoning Research and Development, 50-62. 2005.
2. Berlanga, A.J., López, C., Morales, E. y Peñalvo, F.J. Consideraciones para reforzar el valor de los metadatos en los objetos de aprendizaje (OA). Salamanca. Universidad de Salamanca. Depto. de Informática y Automática. (2005).
3. Chiappe, A., Segovia, Y., Rincon, H.Y.: Toward an instructional design model based on learning objects. Educational Technology Research and Development 55, 671–681 (2007).
4. Chiu, D. (2010). Elasticity in the cloud. ACM Crossroads, 16(3), 3-4.
5. de la Prieta, F., Bajo, J., Marín, P. A. R., & Méndez, N. D. D. (2013). A New Generation of Learning Object Repositories Based on Cloud Computing. In*Management Intelligent Systems* (pp. 99-106). Springer International Publishing.
6. De la Prieta, F., Gil, A. B., Martín, A. J. S., & Zato, C. (2014). Learning Object Repositories with Federated Searcher over the Cloud. In Methodologies and Intelligent Systems for Technology Enhanced Learning (pp. 93-100). Springer International Publishing.Argente, E., Botti, V., & Julian, V. (2011). Gormas: An organizational-oriented methodological guideline for open mas. In Agent-Oriented Software Engineering X (pp. 32-47). Springer Berlin Heidelberg.
7. De la Prieta, F., Rodríguez, S., Bajo, J., & Corchado, J. M. (2013). A Multiagent System for Resource Distribution into a Cloud Computing Environment. In *Advances on Practical Applications of Agents and Multi-Agent Systems* (pp. 37-48). Springer Berlin Heidelberg.

8. Dublin Core Metadata Initiative. DCMI Metadata Terms
9. European Committee for standardization – Cen Workshop Agregament. A simple Query Interface Specification for Learning Repositories. Ref. No.: CWA 15454:2005 E. 2005
10. Fdez-Riverola, F., Corchado, JM. CBR based system for forecasting red tides. Knowledge-Based Systems 16 (5), 321-328. 2003.
11. Frango, I., Omar, N. y Notargiacomo, P. Architecture of Learning Objects Repositories. Learning Objects: standards, metadata, repositories & LMS. Pp. 131-155. (2007)
12. Gil, A. B., De la Prieta, F., & Rodríguez, S. (2011). Automatic Learning Object Extraction and Classification in Heterogeneous Environments. In *Highlights in Practical Applications of Agents and Multiagent Systems* (pp. 109-116). Springer Berlin Heidelberg.
13. Gómez-Sanz, JJ, Pavón, J, Garijo, F (2002). Meta-models for building multi-agent systems. Proceedings of the 2002 ACM symposium on Applied computing, 37-41
14. Hodgins, H. W. (2006). The future of learning objects. Educational Technology-Saddle brook then Englewood cliffs NJ-, 46(1), 49.
15. IEEE Learning Objet Metadata (LOM). Institute of Electrical and Electronics Engineers. http://ltsc.ieee.org. 2002.
16. IMS Digital Repositories Interoperability. Riley, K., McKell, M., y Mason, J. Core Functions Information Model. Version 1.0 Final Specification. (2003)
17. López, V. F., de la Prieta, F., Ogihara, M., & Wong, D. D. (2012). A model for multi-label classification and ranking of learning objects. *Expert Systems with Applications*, *39*(10), 8878-8884.
18. Lujara, S.K., Kissaka, M.M., Bhalaluseca, E.P., Trojer, L.: Learning Objects: A new paradigm for e-learning resource development for secondary schools in Tanzania. World Academy or Science, Engineering and Technology, 102–106 (2007).
19. Open Archives Initiative. Lagoze, C., Van de Sompel, H., Nelson, M., Warner, S. The Open Archives Initiative Protocol for Metadata Harvesting. (2002).
20. Oren, E., Delbru, R., Catasta, M., Cyganiak, R., Stenzhorn, H., & Tummarello, G. (2008). Sindice. com: a document-oriented lookup index for open linked data. *International Journal of Metadata, Semantics and Ontologies*, *3*(1), 37-52.
21. P. Mell and T. Grance. The NIST definition of Cloud Computing. In: NIST Special Publication 800-145. Sep. 2011.
22. SCORM 2004. 4th Edition. March 2009.
23. Simon, B., Massart, D., Van Assche, F., Ternier, S. y Duval, E. Authentication and Session Management. Version 1.0. (2005)
24. von Laszewski, G., Diaz, J., Wang, F., & Fox, G. C. (2012, June). Comparison of multiple cloud frameworks. In Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on (pp. 734-741). IEEE
25. Wen, X., Gu, G., Li, Q., Gao, Y., & Zhang, X. (2012, May). Comparison of open-source cloud management platforms: OpenStack and OpenNebula. InFuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on (pp. 2457-2461). IEEE.