

A New Generation of Learning Object Repositories Based on Cloud Computing

Fernando de la Prieta, Javier Bajo, Paula Andrea Rodríguez Marín,
and Néstor Darío Duque Méndez

Abstract. This work presents a proposal for an architecture based on a cloud computing paradigm that will permit the evolution of current learning resource repositories. This study presents current problems (heterogeneity, interoperability and low performance) of existing repositories, as well as how the proposed model will try to solve them.

Keywords: Learning Object Repository, Cloud Computing, eLearning.

1 Introduction

It has become common in recent years to encapsulate educational resources in the form of learning objects (LO), a process that facilitates their management and reuse. That is, the systematic management of learning resources makes their dissemination possible.

To facilitate these dissemination tasks, the LO are stored in educational repositories. However, these repositories present problems at a technical level, such as low-level performance, unavailability, security, reliability, etc. At the same time,

Fernando de la Prieta

Department of Computer Science and Automation Control, University of Salamanca.

Plaza de la Merced s/n, 37007, Salamanca, Spain

e-mail: fer@usal.es

Javier Bajo

Department of Artificial Intelligence. Technical University of Madrid.

Campus Montegancedo, Boadilla del Monte, Madrid, Spain

e-mail: jbaajo@fi.upm.es

Paula Andrea Rodríguez Marín · Néstor Darío Duque Méndez

National University of Colombia, Colombia

e-mail: {parodriguezma, ndduqueme}@unal.edu.co

the paradigm as a whole also shows many deficiencies like the existence of too many schemas of metadata or interoperability specifications, or even internal architecture of the repositories.

This environment, which is complex not only at the technical level but at the conceptual level as well, has led to a low level of implantation of this technology despite the great advantages derived from its use. Thus, the objective of this study is to propose and subsequently develop a new model which can make it possible to reduce or remove the existing problems. To this end, Cloud Computing paradigm is the key to offering effective and efficient services such as storage, and the search and retrieval of educational resources

The next section presents the state of the art within this context, as well as a study to demonstrate the problems observed in repositories. Section 3 presents the detail of the architecture and, finally, section 4 the conclusions and future work.

2 Learning Objects and Repositories

The encapsulation of education resources in the form of LO makes their reutilization possible. Many authors have recently been presenting their vision regarding this concept [8][4][7], which has led to the appearance of a number of definitions. The aim of this study is not, however, to establish what an LO is, rather to simply remark that there is a clear consensus that an LO must be the minimal reusable unit with a specific learning objective.

There is also a consensus that each LO has to be associated with an external structure of metadata. This metadata allows making a first approach to the educational resource. In other words, the metadata permits improving the utility of the resource, since it makes its retrieval, search, exchange, and hence, its reutilization, possible. The metadata schema is standardized. In fact, there are currently many standards such as DublinCore[3], IEEE LOM[8], etc. The existence of standards facilitates the management of the resources, enabling the interoperability among systems that use compatible standards.

Although at first sight these standards can be seen as an advantage, reality shows that in some cases they are the problem, as many existing standards are not compatible among themselves. The ADLNet¹ 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 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. To this end, 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. [2].

As with traditional education resources, LOs are stored in libraries, in this case digital libraries called repositories. A digital repository can be defined as a place

¹ Advanced Distributed Learning Network. <http://www.adlnet.org/>

where a digital resource can be stored, searched and retrieved. These systems must also support the import, export, identification and retrieval of content. [11]. The LOs are usually stored in a Learning Object Repository (LOR). JORUM project [10] states that an LOR is a set of LOs with detailed and external information (metadata) that is accessible from the Internet. In addition to housing the metadata, an LOR can also store the educational resource. In general terms, an LOR must implement the following task [**Error! No se encuentra el origen de la referencia.**]: Search/Find, Ask, Retrieve, Send, Store, Collect and Publish.

The deployment infrastructure can basically be either distributed or centralized. Taking into account that an LO is formed by a digital resource and its metadata, there are four kinds of possible infrastructures [6]: (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. Furthermore, three kinds of storage strategies can be distinguished [6]: (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.

A controversial aspect is the interoperability among LORs. Firstly, it is necessary for a repository to use at least one standard in the stored metadata. And secondly, there must be an interface from where an external search agent (a client or another LOR) can access the stored information. This interface is currently implemented through high-level interoperability layers that allow access from outside the LOR. There are different standards or specifications that focuses on this interoperability layer:

- **OAI-MPH** (*The Open Archives Initiative Protocol for Metadata harvesting*)[12]. This protocol provides a technology-independent framework for retrieving documents or resources, thus enabling interoperability among systems. The protocol is open and the repository is not limited to educational resources.
- **IMS DRI** (*IMS Digital Repository Interoperability*) [9]. This is based on existing communication technologies and on previous specifications of the IMS consortium. This protocol is immature and is still in its initial stages of development.
- **SQI** (*Simple Query Interface*) [4]. The kernel of SQI 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 interfaces supports synchronous/asynchronous and stateful/stateless queries. The authentication is based on a session with the aim of isolating the harvesting of contents from the management tasks.

The use of an abstraction layer between LOR and the client system avoids the need to take the internal infrastructure of the repository into account, and allows clients to perform queries to many repositories in parallel. In other words, the clients can perform federated searches [1][5].

2.1 The Reality of LOR Interoperability

The state of the art shows a high heterogeneity in existing standards. Therefore, a study of LOR has been performed in order to analyze the real situation. The study includes the analysis of the following LORs: *Acknowledge*, *Agrega*, *Ariadne*, *AriadneNext*, *CGIAR*, *EducaNext*, *LACLO-FLOR*, *LORNET*, *MACE*, *Merlot*, *Nime*, *OER Commons* and *Edna Online*. It consists of performing 60 queries to each LOR through an SQI layer that the repositories provide. The query patterns are Unesco codes.

Firstly, the general characteristics of each LOR are analyzed; all of them use IEEE LOM as metadata schema and VSQL [15] as query language. Additionally, the majority of them are stateless (65%), and all of them have synchronous interfaces, but only 4 have the asynchronous interface.

Considering that the SQI specification does not force the implementation of all methods of the specification, the SQI methods that these repositories have implemented are checked. The results radically change the outlook because 6 of the 14 repositories do not work or are unavailable and they have to be removed from the scope of this study (*Ariadne*, *AriadneNext*, *EducaNext*, *Nime* and *EdNa Online*). *MACE* and *LOCLO-FLOR* produce an error in the authentication, although the process is done correctly. After this step, this test is 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.

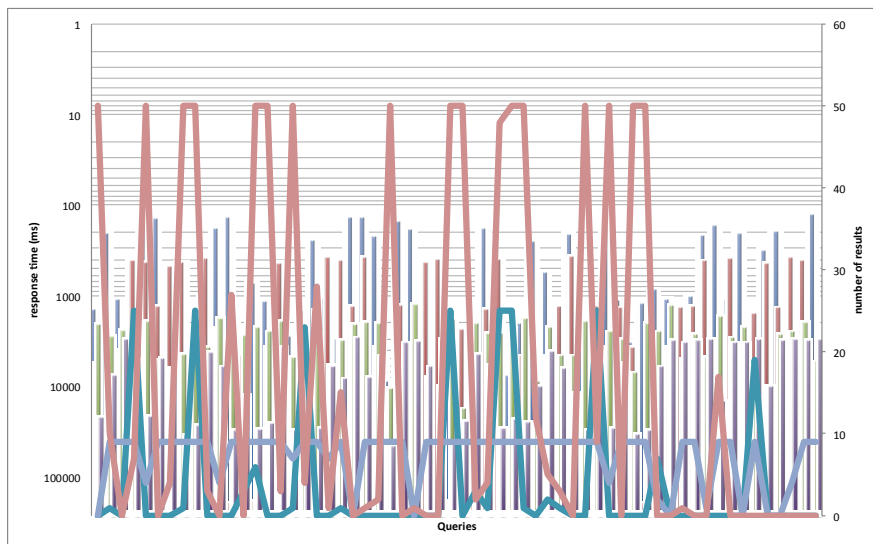


Fig. 1 Repositories performance

Finally, to summarize this study, Figure 1 shows the performance of each LOR. The lines are the number of results and the bars are the response times. The stateful repositories are slower than stateless repositories, with an average 19,1685 seconds compared to 3.053 seconds respectively, because they have to authenticate the client before performing the query. The average of the results is 9.56 LOs retrieved per query.

3 An Opportunity Focused on Cloud Computing

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.

Lately, within the services in the context of Internet, Cloud Computing is emerging as key paradigm of the present century. According to NIST² [13], *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).

However, beyond the kinds of services, the key characteristic of this new paradigm is the quality of services. 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.

Taking into account the weakness that has been demonstrated in this study with regard to the performance, availability and interoperability of existing LORs, this study proposes a new architecture based on Cloud Computing.

This new architecture will make use of the services that +Cloud platform [14] provides, such as storage and databases. This platform is based on the Cloud Computing paradigm. This platform allows offering services at the PaaS and SaaS levels. The IaaS layer is composed of a physical environment that allows the abstraction of resources into virtual machines.

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 *IdentityManager*, which is the module of +Cloud in charge of offering authentication services to clients and applications.
- the *File Storage Service (FSS)*, which provides an interface for a container of files, emulating a directory structure in which the files are stored with a set of metadata, thus facilitating retrieval, indexing, search, etc.

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

- the *Object Storage Service (OSS)*, which provides a simple and flexible schemaless data base service oriented towards documents.

3.1 Proposal: CLOR

This study proposes the development of a new platform called CLOR (Cloud-based Learning Object Repository) based on +Cloud as its underlying architecture. Figure 2 shows a diagram with the main components of this modern architecture. The details of each component are presented as follows:

- **CLOR Management** is the kernel of the architecture. 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; at the same time, it is in charge of providing a set of functionalities in terms of web services to the upper layers, that is, to the end user interfaces.
 - FSS will be used to store the educational resources. FSS also encapsulates the traditional complexity of the file system storage; this component only has to call web services in order to retrieve or store resources. Furthermore, because of other FSS functionalities, such as file versions, metadata associated with each resource, etc., it will be possible to increase the power of the service. Finally, it should be noted that the elasticity of the FSS implies no limitation regarding storage capacity.
 - OSS will be used to store the metadata associated with each learning resource. OSS makes use of a nonSQL database that permits storing the metadata en JSON format. The main advantage is that it permits storing any kind of metadata independent of its structure or schema, that is, its standard. Furthermore, queries about the LO will be performed very quickly thanks to the underlying database.

This component will be complemented with different interoperability layers, such as SFI or OAI-MPH, which will ensure the communication with other LORs and federated searches from external clients.

- **CLOR GUI** is the graphical user interface of the end users. The key architectural characteristic is that it will be independent of the bottom layer (in other words, the CLOR Management) and communication will be carried out through web services. The management of the users will be delegated to the Cloud, specifically, to the identity manager.

CLOR will present two independent interfaces:

- *Storage CLOR* in charge of managing the repository (storage and creation of metadata).
- *Search CLOR* will provide an interface to perform queries not only in this proposed repository, but also in the resources of other resources by means of the interoperability layers. This interface component will, in the future, also provide other functionalities such as a recommendation system, a key functionality that has recently emerged [1].

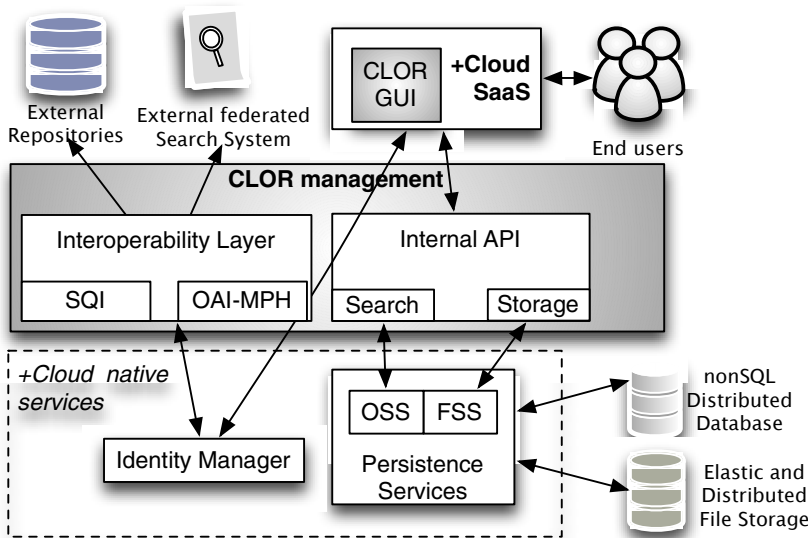


Fig. 2 CLOR architecture

4 Discussion, Advantages and Future Work

This study has presented an innovative architecture that constitutes an evolution over current storage system for educational resources. This new model, will enable the observed problems, which have been demonstrated in this study, to be solved:

- *High heterogeneity in terms of number and characteristics of existing standards.* The proposed model allows dealing with the heterogeneity of current and future standards since it is based on a non-relational database.
- *Low performance.* Cloud computing paradigm allows offering services with the same level of quality independently of its demand. The development of the LOR based on this paradigm will make it possible not only to offer an effective service effective, but to offer an unlimited storage capacity as well.
- *Interoperability among repositories.* The low linkage among components permits implementing many interoperability layers without needing to upgrade to other modules.
- *Complementary services.* This model will make it possible to include other functionalities in its own repository that until now were not possible, such as recommendation model, space of storage for each user in the cloud, a collaborative model for creating learning resources and metadata, etc.

Future work will be focused on finishing the development of the proposed repository and evaluating it not only at a technical level, but also at a functional level by using end users that are teachers and students.

References

1. Gil, A.B., Rodríguez, S., de la Prieta, F., De Paz, J.F.: Personalization on E-Content Retrieval Based on Semantic Web Services. *International Journal of Computer Information Systems and Industrial Management Applications* 5(2013), 243–251 (2010) ISSN 2150-7988
2. Berlanga, A.J., López, C., Morales, E., 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. Dublin Core Metadata Initiative. DCMI Metadata Terms
4. European Committee for standardization – Cen Workshop Agregament. A simple Query Interface Specification for Learning Repositories. Ref. No.: CWA 15454:2005 E (2005)
5. De la Prieta, F., Gil, A.B., Rodríguez, S., Martín, B.: BRENHET2, A MAS to Facilitate the Reutilization of LOs through Federated Search. *Trends in Practical Applications of Agents and Multiagent Systems*, 177–184
6. Frango, I., Omar, N., Notargiacomo, P.: Architecture of Learning Objects Repositories. *Learning Objects: standards, metadata, repositories & LMS*, pp. 131–155 (2007)
7. Frango, I., Omar, N.: Architecture of Learning Objects Repositories. *Learning Objects: standards, metadata, repositories & LMS*, 131–155 (2007)
8. IEEE Learning Object Metadata (LOM). Institute of Electrical and Electronics Engineers (2002), <http://ltsc.ieee.org>
9. IMS Digital Repositories Interoperability. Riley, K., McKell, M., y Mason, J. Core Functions Information Model. Version 1.0 Final Specification (2003)
10. JISC Online Repository for [learning and teaching] Materials (2004)
11. Joint Information System Committee. What is a Digital Repository? Digital Repositories. Helping Universities And Colleges (2010)
12. Lagoze, C., Van de Sompel, H., Nelson, M., Warner, S.: The Open Archives Initiative Protocol for Metadata Harvesting (2002)
13. Mell, P., Grance, T.: The NIST definition of Cloud Computing. In: NIST Special Publication 800-145 (September 2011)
14. Heras, S., De la Prieta, F., Julian, V., Rodríguez, S., Botti, V., Bajo, J., Corchado, J.M.: Agreement technologies and their use in cloud computing environments. *Progress in Artificial Intelligence* 1(4) (2012)
15. Simon, B., Massart, D., Van Assche, F., Ternier, S., Duval, E.: Authentication and Session Management. Version 1.0. (2005)