FARE: A New Free Architecture for Remote Education

Leonardo Favario, Angelo Raffaele Meo, Enrico Masala
Dipartimento di Automatica e Informatica – Politecnico di Torino
Corso Duca degli Abruzzi, 24 – 10129 Torino Italy
leonardo.favario@polito.it, meo@polito.it, masala@polito.it

In recent years the amount of educational resources available online has constantly increased. However, it remains challenging how to make use of such wealth of material in a simple and effective way, ranging from primary schools to high schools. Typically, the lack of available time and the peculiar procedures imposed by each system constitute the main difficulties faced by the teachers. In this context, this work proposes an architecture to allow different repository systems to work together in a transparent fashion by means of the CMIS API and through a new integration layer with a simplified interface, designed on the basis of the needs of typical content creators and users. The preliminary feedbacks are encouraging, showing that the platform could be widespread adopted by the Italian educational system.

1. Introduction

A large amount of tools are currently available in the e-learning environment, therefore educators have a vast range of possibilities to innovate their teaching styles. Many of these tools include software that allow to store learning objects (LOs) in large repositories and to properly modify and reorganize them into lectures suitable for the teachers’ needs. The tools also include Enterprise Content Management (ECM) systems such as Alfresco, Nuxeo, Sharepoint, etc. More specific tools are available for educational purposes. These are the so called Learning Object Repositories (LOR), which include, among others, MERLOT, MIT OCW, Ariadne. If learning content creation features are also present, they are generally referred to as Learning Management Systems (LMS). Examples are Moodle, Edmodo, Blackboard.

In the past, teachers have created a large number of contents in digital format, that were made available typically using manual procedures, without an integrated content management system, on personal websites. Clearly, their format strongly varies due to the use of different software to create them. For this reason, currently LO repositories (LOR) allow to store different types of
resources. However, differently from the past, they are made available in a more ordered manner by means of indexing and tagging operations.

Though there are a number of international standards to format LOs, there might be many reasons for the content creators not to follow such standards (e.g., lack of appropriate tools, familiarity with the tool themselves, etc.). Therefore, these objects are typically organized in a repository at each institution that allows on-line access to the learning material, by means of systems that, in the majority of cases, are not the same between similar institutions. That is due to a number of factors, including past IT choices, availability of trained staff, etc. Moreover, the choice often happened without involving the teachers, or even worse, asking them to use the new tools only on a voluntary basis as a complementary activity. Therefore, there is a significant risk that many teachers that were initially motivated in creating online contents are now not sufficiently motivated to use the new technology.

However, recently there is a strong momentum in going towards integration of repositories of different institutions, often backed up by the national bodies such as the Ministry of Education, as it is currently happening in Italy. Integration, in general, poses many problems, from both the technical and the social point of view. For the former case, it is difficult to have common schemes for metadata so that content can be easily searched and used in an integrated manner across repositories. Concerning the latter case, teachers have a strong desire to adapt the tools to their peculiar need, that is typically widely different depending on the level and even on the teaching style itself. Although we refer to the specific Italian case, we deem that many of the issues encountered here might be encountered also in other national contexts.

A possibility to overcome the previous difficulties, at least partially, could be to find ways to present contents to the teachers, creators and integrators in a simple and effective way through appropriate user interfaces. Moreover, there is a strong need of integrating the large amount of material that is currently available online, but in many different repositories, by means of facilitating the creation of direct links between the various objects regardless of their location. This can be done, in our opinion, by means of a single point of access, though the functionalities have to be limited due to the different features available in the various CMS or LMS.

The addition of metadata to content may facilitate such operation. In fact, this has been done also by relatively old projects such as MERLOT [Burns and Schell, 2002], where academic institutions collaborated to create a homogeneous repository that was easily indexed and searched by a central location. Moreover, LORs can exploit the metadata to create the so-called learning paths. Many standards are available to define the semantic of the metadata information added [Ochoa et al, 2014]. In addition, tags and other comment fields may be included to enrich the content definition.

An important point to consider is the openness of the content and the software itself. In fact, having Open Educational Resources [Wenk, 2010]
facilities diffusion and usage of the LOs. Although a formal definition of an open content as the one in [Hylen, 2006] is not strictly needed, this however facilitates the reuse, remix and modification by removing any potential ambiguities.

Once legal issues are cleared, technical solutions are needed to share LOs between different institutions using a single access point. The current trend is to use Unified E-learning Repositories (UER). In this context, a number of works tried to solve the numerous integration issues that happen in practice. The GLOBE [GLOBE, 2015] initiative allowed clients connected to a central point to run queries that were automatically propagated to all appropriately configured repositories in the network [Ternier et al, 2008]. Lower-layer technical solutions also need to be solved, as pointed out in [Hatala et al, 2007]. If repositories are homogeneous, the integration task is greatly simplified, as in the case of CampusConnect [Boehringer and Bernlohr, 2014]. Other systems, such as the one presented in [LON-CAPA],[LIONSHARE], and [POOL] tried to address the case of heterogeneous environments, however many constraints are still present, which do not facilitate the adoption of such solutions.

In this context, we propose to use the CMIS interface [Choy et at, 2010] that provides a standardized API for the most common operations. This interface is supported by several systems, including the widespread Alfresco and Nuxeo ones. Although the set of functionalities is more limited with respect to the case of homogeneous environments, we argue that this interface is sufficient to build a successful service that can adequately address the needs of most of the teachers.

On the basis of several studies that have been conducted by means of interviewing teachers about their preferences in the context of e-learning systems, we defined a set of requirements for our project, which are reported in the following:

1. **Free software**
   This is important so that the students have both the tools and the materials for learning with full accessibility and no additional or hidden costs. The software has to be released using a GPL based license, which also allows to be compliant with the Italian regulations on Public Administration software adoption.

2. **All materials available in Creative Commons**
   Since a very important operation in creating teaching material is the reuse of part of existing material, this type of licensing is required to ensure that the material used as a source is available for this purpose. We selected the Creative Commons BY-SA 3.0 IT license as the default one for any new object uploaded to the repository.

3. **Simple interface** for access and usability
   We desired that the usage experience was as much immediate and intuitive as possible, without the need to learn new tools or particular procedures.
4. Possibility of running **optimized queries**
The system should be capable of running queries, a fundamental operation to locate the content of interest, in the most efficient way, considering that objects can be located in physically different repositories and systems.

5. **Scalability** and possibility of **diffusion**
The scalability factor becomes important as the solution will be adopted, as we hope, by a multitude of students and teachers, while the diffusion is implicitly guaranteed by the availability of the source code.

The rest of the paper presents our solution to the previous problems, including the discussion of the results achieved by a pilot phase that involved some teaching material made available by the two universities of our city. However, the solution aims at including also schools at different levels and from different regions. In this context, this project has also been proposed for consideration at the Italian Ministry of Education as one of the possible platforms for a national initiative in online learning for Italian schools.

2. **FARE – The Free Architecture for Remote Education**

The purpose of this project is the set up of a network of interconnected repositories. By means of a series of inter-operable tools, the users will be able to define their personalized learning paths leveraging the vast variety of materials available. The FARE platform, whose acronym stands for Free Architecture for Remote Education, is based on access freedom, both for the source code and the contents.

2.1 **Overview**

As mentioned above, thanks to this integration the final users will be able to transparently access a set of resources which can be geographically distributed in different places. The distributed design and its modularity make it possible to add several other repository nodes without particular efforts.

The main entities involved in the architecture are summarized below:

1. **LAMP environment**
   For the sake of building the prototype a LAMP environment was deployed. This implies the provision of a GNU/Linux operating system installed on a server with the Apache Server responsible of serving HTTP requests, a MySQL relational database management system and PHP as a scripting language.

2. **Web application**
   Drupal 7.x was selected as a framework for the web application because of its modularity, scalability, user management features and the availability of a strong community that actively supports and maintains all the modules. The Drupal framework runs in the LAMP environment, taking advantage of both the PHP and the MySQL database. The Bootstrap framework handles the frontend of the web application.
3. **Repository**

The application can be connected to several repositories completely different from each other thanks to the use of CMIS. For our benchmarks we selected Alfresco Community Edition (CE) and Nuxeo but other repository software solutions have been proved working (i.e. Sharepoint). Our selection was based on the level of CMIS mapping, the community support and the market share in the educational environment. Since Drupal is mainly written using PHP, in the development of the connector module the **CMIS PHP API** was used.

4. **CMIS (Content Management Interoperability Services)** is a standard for interoperability between ECMs. This standard allows to connect together repositories of different nature without using a custom-written proxy to translate each call in different proprietary languages. Using the API it is possible to achieve a high level of flexibility, writing code just once for all the CMIS compliant repositories. The hierarchical structure of FARE is perfectly aligned with the CMIS model since both are based on two entities: folders and documents.

5. **Learning object**

The platform can handle an extensive variety of educational resources. This objects can be classified as simple (i.e. text files and pictures) or complex such as SCORM files. Each LO is enriched by proper metadata whose values are stored in the database of the repository where the LO resides.

The web application offers a single web interface for both teachers and learners. Once logged in, the user can reach different tools through its own personal dashboard. In particular it is possible to perform numerous actions which can be summarized as upload, search, selection and results organization.

The upload phase consists on loading new contents in the repository. Web forms will help the user to fill the metadata information which are crucial for searching. During the search phase the query is formulated through the web application, translated in **CMIS SQL** language and sent to the CMIS interfaces of all repositories. This federated model has been proved fitting better with a heterogeneous environment than the indexed one and the performance drawbacks registered are negligible. Once the search returns the results it is possible to explore the desired ones, save the entire selection or perform different actions on the selected contents. In details, contents can be exported in a compressed format, merged together, cut in parts and pasted in a new file. Lastly the document can be downloaded, shared or exported.

2.2 **Use cases**

To provide a clear explanation of the possibilities offered by the platform two use cases are presented below.

1) Teachers: The homepage offers the possibility to search for topics. In the results table the desired contents can be selected and immediately exported in
ZIP format or some further editing can be carried out. The compressed file contains the selected material as well as other supporting files to allow an easy use and presentation of the material. More precisely, two HTML5 slide-shows are included in the archive. The first contains the meta-information about the content itself whilst the second one allows to display the content easily even on an interactive whiteboard environment. This has been achieved by depicting big icons to ease the selection also with imprecise pointing devices (e.g., using interactive whiteboards). HTML5 standards have been selected to guarantee the possibility to easily display the content with the same style on different platforms.

After the selection process, the teacher can benefit from a set of editing tools. It is possible to reorder the contents, to select just some parts of it and to export the aggregate selection as a single PDF file. This is very convenient for teachers preparing content by splicing together several documents (i.e. picking only some pages of a resource) since it allows to share a single complex LO instead of a collection of links to sparse contents. The final document will be automatically created on-the-fly and can exported.

During the upload phase, an authenticated user with upload permissions (typically a teacher) can populate the repository with new contents providing also an appropriate metadata definition. When a complex LO is created, comments can also be added so that other teachers could benefit from those extra side information. This again leads to a new set of metadata which can be useful in the following researches.

2) Learners: Learners typically use the web application for two reasons: they received a URL from a teacher pointing to some resources already prepared for them or they want to perform a new search. In the first case the URL shows not only the final document as edited by the author but it also provides direct pointers to all the sources used in the document. Thanks to this flexible solution, the learners can decide to download the document, to check which sources have been used and to download only a subset of them. This is particularly convenient in the case where the student is accessing the platform with different devices since it is possible to choose the download option which best fits the current needs.

2.3 First experiments

The described system has been implemented and it is currently tested at our university in a live environment. At the time of writing two repositories are involved, i.e., the Alfresco repository residing in our University and a Nuxeo repository used for tests, that will soon be replaced by the Nuxeo repository of the other university in our city.

There are a few registered users that performed and tested several actions, including document upload, reaching 700 heterogeneous learning objects including text documents, images, video lectures and covering 18 subjects. A total of 15 complex LOs have been prepared by users using the cut and merge feature drawing from both repositories.
Usually the complexity introduced by a distributed architecture causes performance loss. Several performance tests have been carried out to understand during which of the normal operations the platform can suffer some drawbacks. The first tests measured the response time of a general query run on different distributed repositories. The results are encouraging as the latency for such a broad scope query is, at maximum, about 500 ms. Since the query will be performed by means of the web application, the overall overhead introduced by the web stack has been tested as well. This results in about 400ms overhead but can be slightly improved by means of some interface optimization tricks, such as aggregation and compression of CSS and JavaScript resources.

As above mentioned, one of the most innovative feature is the possibility of mixing the resources on-the-fly and finally merge them into a single resulting file. By means of automated scripts also these functionalities have been tested and the results can be deemed acceptable as they are in the order of 1-2 seconds.

When creating a final ZIP archive, the system has to retrieve all the documents, compress them and finally prompt the user with a download window. Taking in account the distributed nature, the timing of the retrieval operation can vary since the location, the infrastructure and the different configuration of each repository platform entail an increase in latency. Alfresco performed faster than Nuxeo and this could be due to the authentication process that is slower on the latter.

We are already testing a second version of the platform which builds on the suggestions of some early users of the system. In particular, we included a set of HTML5 templates to simplify the presentation of the content. The most critical issue, that is, the responsiveness of the web application, did not raise concerns by any of the users at this time. Therefore, the performance presented in this result section can be considered acceptable for the aforementioned purposes.

3. Conclusion

As mentioned in the introduction, an increasing number of learning objects is available on a plentiful of online platforms. The absence of an integration between most of the architectures results in practical problems for the final users which become discouraged and abandon the e-learning approach. The main focus of the project FARE is to tackle this situation by offering an easy-to-use interface that transparently integrates several different technologies together.

The friendliness and usability of the interface was confirmed by the first user feedbacks. This can lead to a significant impact on the quantity of LOs in Italian language available on the Internet which is crucial to improve the remote learning perspective.
The possibility to easily integrate different repositories together could potentially allow not only schools but also other cultural institutions to share their material. Educators could also benefit from this since the number of learning resources available in Italian is nowadays too low. Hopefully, in this context FARE will become an application that can provide the features of both a search engine and an LMS, allowing to search for objects that can be further elaborated without using any other external specific tool. This perspective is particularly interesting for its adaptability to widely different environments beyond the educational one.

Bibliography


