Managing Open Educational Resources on the Web of Data

Presenting COMETE, an RDF-based Resource Manager

Gilbert Paquette CICE Research Chair, LICEF Research Center Télé-université Montreal, Canada

Abstract-In the last few years, the international work on Massive Open On-line Courses (MOOCS) underlined new needs for open educational resources (OER) management within the context of the Web of Data. First, within MOOCs, all (or at least most) resources must be open and available on the Web through URIs, including the MOOCs themselves. Second the evolution of research and practice in the field of OER repositories, notably the focus in international e-learning standards, is moving recently from OER metadata stored in relational databases towards RDF-based descriptions of resources stored in triple stores. Third, new resource management tools like COMETE provide more intelligent search capabilities within the Web of data, both for designers who are building MOOCs, and also for students who should be equipped with friendly tools to personalize their environment. We will present some COMETE use cases to illustrate these new possibilities and advocate for their integration within MOOC platforms.

Keywords—resource management; resource repositories; OER; open resources; MOOCS; RDF; Web of data; Semantic Web; IEEE-LOM; ISO-MLR

I. INTRODUCTION

This paper results from a decade of research started in 2002-2004, within the eduSource canadian project, where a resource manager called PALOMA [1] was first produced and improved in the following years. Later on, the main research moved within the Learning Object Repository Network (LORNET), a large pan-canadian 5-year effort directed by the first author. This effort resulted in the production of TELOS [2], a design workbench for learning scenario or knowledge management workflows based on a new manager describing resources in the terms of an OWL technical ontology. In the last three years, our research on educational resource management moved to the use of semantic technologies for the Web of data [3,4,5], resulting in a mature tool, COMETE, that is being used in the colleges of Quebec for educational resource referencing and search.

The present paper summarizes this recent effort for the first time, describing the COMETE system and linking it to the ISO-MLR standard [6], proposing its use as a tool for the design and use of massive open on-line courses (MOOCs) [7,8].

The paper in organized into four more sections. In section 2, we introduce the notion of an educational resource repository and of a resource manager. Section 3 present a

Alexis Miara CICE Research Chair, LICEF Research Center Télé-université Montreal, Canada

recent evolution of e-learning standards, norms and application profiles resulting in the publication of the ISO standard on Metadata for Learning Resources (ISO-MLR), which is based on the Ressource Description Framework (RDF) enabling the use of the Web of data. In the fourth section we provide an overview and some insights on COMETE, our RDF-based OER manager. In the last section, we use COMETE to illustrate its use both for MOOC design by professors or designers, and for MOOC personalization by students and tutors.

II. OPEN EDUCATIONAL RESOURCES REPOSITORIES

We provide here some backround information on open educational resources repositories and resource management.

A. Open educational resources repositories

The term "Open educational resources" was first coined at UNESCO's 2002 Forum on Open Courseware and defined as "teaching, learning and research materials in any medium, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions. Open licensing is built within the existing framework of intellectual property rights as defined by relevant international conventions and respects the authorship of the work."[9]

Ten year after, UNESCO held in Paris an international OER congress on 20-22 June 2012 where the so-called Paris OER Declaration was issued. This declaration recommends that States, within their capacities and authority "foster awareness and use of OER" in many ways including "encourage research on OER", "promote the understanding and use of open licensing frameworks", and "facilitate finding, retrieving and sharing of OER."

This last recommendation refers to the important on-going work in the last ten years on so-called "learning objects repositories (LOR)". Many definitions have been given for "learning objects (LO)". It is a more general concept than OER, since not all LOs are open, some being copyright protected. But still, LO repositories use the same methods and technologies than OER repositories.

B. First interoperability norms: Dublin Core and IEEE-LOM

The idea that educational contents could be seen as "objects" to be reused in multiple contexts dates back to the

late 60's but it started to become a reality only by the middle of the 90s with the generalization of the Internet [10].

In 1995, an international consensus arose around the necessity of e-learning standards to promote tools' interoperability and learning objects reusability. The aim was to insure the reuse of educational objects jeopardized by the diversity of referencing metadata schema around the world. This goal was shortly concretized by the Dublin Core (DC) metadata initiative [11] proposing a first set of standardized metadata, expressed in XML. Since then, the Dublin Core metadata schema has become one of the most used vocabularies on the Web of Data.

In 1996, the IEEE created the Learning Technology Standards Committee to integrate previous work on the concept of Learning Object Metadata (LOM) [12]. In June 2002, on the basis of a joint IMS-ARIADNE proposal, IEEE approved a LOM standard that was largely accepted internationally. From then on, major resource repository initiatives bloomed rapidly: ARIADNE in Europe, MERLOT in the USA, EdNA in Australia.. These and many other organizations, including our own LORNET, joined the GLOBE [13] consortium that operates actually a large repository of nearly one million resources. A metadata record is a standardized set of properties of a learning object that makes its retrieval possible throughout the world using computer software as if the resources were in a unique reservoir, whatever their actual location.

A resource manager is a piece of software that provides the essential functionalities to make it work. Fig. 1 presents a view of the PALOMA resource manager [1]. Most LOM-based managers like this one have the following components.

- The *Harvester* or *Metadata Repository Builder* help find the location of interesting LOs on the Web or on a Local area network and creates a LOM record for each LO or resource. To make a resource more widely available, this component will sometimes transport it on some predefined server location. Harvesting is now the main method to add metadata records in a repository. It rests on the OAI-PMH protocol [14], based on a client–server architecture, in which "harvesters" request information on updated records from repositories. In this way metadata records already present in some repositories can be selected and grouped in larger or more specialized repositories.
- The *Metadata Editor* provides forms to enter all the metadata for any resource and stores the metadata

| 000 | Paloma | |
|---|---|---|
| Settings Profiles Repositories Help | | |
| Folders 🔑 Search | | 📄 Metatagger |
| Banque GestionCC Banque IDLD Banque IDLD Schartios GTN-Québec ScORM ScoRM Site Carrefour Éducation Site PerfecTIC Site Saut Quantique Add Delete Title Repository Local Attachage de tiges (Ic Banque LD-PRIOWS http: Mathématiques pour les c Banque LD-PRIOWS http: Mathématiques pour les c Banque LD-PRIOWS http: | Profile : IEEE LOM 3 - Met 4 - Tec 5 - Edu 6 - Rig 7 Learning Design 2 Source Classifica Design d'apr A 510-Na Ressource A131-Gr Ressource A135-Im Ressource A135-I | Rel 8 - Ann Ition Irrative Text aph/Diagram age: Photo/Drawing/P imputer-based Unime istry bre métrie uation différentielle Équation aux dérivées Matrice 2-Déterminant Espace vectoriel -Série de Fourier -Méthode d'intégration |
| | Add Remove | Display Path |
| 1 element selected | Light view Standard view | |
| 🛨 Add 🖃 Delete | 🗖 Launch 📄 HTML 🔮 XML | 🔙 Save |

Fig. 1. Example of an early OER manager: PALOMA [1]

C. Resources managers based on Learning Object Metatda

The learning object (or OER) repositories are basically computer servers, databases and management software operating on the Web that can deliver the learning objects to any computer connected to the Internet. The learning objects are described by metadata stored in databases. In most configurations, the metadata is separated from the resources that can reside on one or more servers connected to the Web. record in a permanent relational/XML database. Fig.1 shows on the right part some metadata entered for the selected resource within the selected repository/folder on the left. The Metadata shown here is selected in section 9 of LOM standard where terms in classifications can be chosen.

• The *Repository Search Agents* apply user-defined constraints involving metadata properties to find and display a set of the corresponding metadata records for a user to select and view the metadata and the resource.

- The Access Manager helps define a user's profile and its access rights to folders, metadata records and resources. User rights may include viewing, adding, modifying, deleting and assigning rights to other users.
- The *Repository Structure Manager* is a commodity for grouping resources' metadata records into folders. It can move a metadata record from one folder to the other, copy an alias in another folder, suppress a record or a folder, and duplicate a record to speed up the metatagging of a similar resource.
- The *Collaborative Annotator* provides message editing by users about a resource. It also provide ratings functionnalities that can be consulted by other users, sometimes offering resource display according to these ratings.

D. Potential and Limits of DC/LOM Resource Repositories

There is an large interest around the world for learning object or resource repositories as exemplified by the number of existing repositories, of organizations building and sustaining them, of contributors integrating learning objects in repositories and of the users of these learning objects. Most international conferences and journals on technology-based learning include scientific communications, some journals being specifically devoted to the subject¹.

The fundamental reasons for this interest are the growing educational demands in all countries, the limited capacity of face to face education to fulfill the demand in a timely manner, the important effort and cost involved to build online multimedia learning materials and the new possibilities offered by the Internet.

While it is a fact that millions of documents can be found on the Internet using search engines like Google, there is no guarantee that a query will lead to trustable material on which high quality education can be built. Learning object repositories offer a solution to this problem.

- First, resources repositories are maintained by educational institutions and professors that put their expertise and credibility in the balance, providing a certain trust in the quality of the referenced resources.
- Second, these resources are often peer-reviewed to ensure their quality; comments on the documents are made on the repository website to identify their actual use and their reusability capacity in different areas.
- Third, the metadata associated to the learning objects give precious information to the users, such as the name and location of the authors, the type of learning or teaching resource, the knowledge contained in it, the educational use that can be made of it, the languages in which it is delivered and the technical requirements for its proper use.
- Fourth, this metadata serves to make focused queries according to a user needs based on the properties of the

resource, not only on vague keywords that leads to thousands of references that one needs to read to understand what kind of content they provide.

• Finally, the vast majority of these learning objects are in the public domain. They are OER to be reused free of charge. The resources can be adapted or aggregated, and referenced back in a repository to extend the availability of good learning material.

After a decade of research and practice in this field, although they provide a solution to cope with the growing educational needs of the knowledge society, there are still a number of limitations to a larger use of OER repositories.

- *Cultural issues, author recognition.* Many authors will keep for themselves and their students the resources they build for their courses. While research papers are easily shared, educational resources are seen by many as private property that should be protected by copyright. Some repository like Merlot provide various kind of recognition and rewards for authors who share openly, but these methods not as widespread as in the case of research papers.
- *Rigid, closed institutional systems.* Many educational institutions keep their resources in house, integrated closely within online course stored in a learning content management system (LCMS) or online course platform, thus preventing a larger use of the ressources.
- *Slow standard adoption.* Without the adoption of an international standard, resources described by proprietary metadata remain local, unshared from one institution to another.
- *Multiplicity of IP holders.* Another severe drawback is the slow adoption of open licences like Creative Commons or GPL/LGPL, blocking the reusability of learning ressources, moreover if the resource is complex and subject to multiple intellectual property (IP) rights.
- *Heavy referencing process.* Even when the LOM standard is adopted and resources are submitted to open licences, the referencing process is complex. The LOM has 9 sections and 86 possible metadata entries so most institution will reduce the metadata set to a LOM profile covering only part of the standard and they will add specific vocabularies, thus facilitating the referencing process, but complicating search operation of resources coming from multiple repositories.

III. ISO-MLR AND THE WEB OF DATA

Many of these limitations will be overcome through more information to institutions and authors on standards and open licences. Still others require new approaches such as the Web of linked data.

A. ISO-MLR: an OER referencing standard based on RDF

Although the Dublin Cores and the IEEE-LOM are widely used to describe learning resources, interoperability among

¹ Please consult the IJKLO journal at www.ijklo.org and its successor at http://www.informingscience.us/icarus/journals/ijello

metadata sets from multiple repositories is still challenging, as best practices are only recommended.

For example, instead of using ISO 8601, a DC Date element can be written in plain language making impossible its processing by queries. Ambiguous definitions pose another challenge. For example a Date element can represent a resource creation time, a time of update or a time of publication. As mentionned above, LOM records can be based on a wide variety of Application profiles each defined in their own way by various agencies.

The ISO/IEC 19788 standard [7], in short ISO-MLR, is intended to provide optimal compatibility with both DC and the LOM. It presents the following advantages.

- Insuring the coherence and the non-duplication of concepts by proposing an RDF-based data model.
- Preventing the proliferation of non interoperable application profiles.
- Supporting the extension of description vocabulairies in precise ways while preserving interopérability.
- Supporting multilingual and cultural adaptability requirements from a global perspective.
- Integrating ressource referencing and search with other data sets in the Web of linked data.

The graph in Fig.2 shows part of the ISO-MLR RDF model. The ovals represent classes of resources, the rectangles are value types, properties are written on the links. This graph summarizes the RDF triples in the section 5 of the standard. Here are some of the triples present on Fig.2:

(Learning resource, has learning activity, Learning activity) (Learning activity, learning method, *method value*) (Learning resource, has contribution, Contribution) (Contribution, has contributor, Person) (Learning resource, has annotation, Annotation) (Annotation, annotation date, *date value*) (Annotation , annotator, Person)



Fig. 2. Part of the ISO-MLR RDF model

B. A standard for the Web of data

The fundamental thing here is that ISO-MLR proceeds from a different vision than previous standards like the IEEE-LOM, where resources as seen only as documents. ISO-MLR, using technologies like RDF and RDF schema, integrates well to a Web of linked data, instead of simply a Web of documents.

The origin of the Web of data, also termed "Semantic Web", dates back to 2001 when the actual director and founder of the Web, Tim Berners-Lee and his colleagues [15] proposed to integrate to Web pages information on the knowledge (concepts, properties) present in Web documents. The URLs who provide locations on the Web were to be generalized to URIs that could represent people, real-world objects, but also abstract concept and properties. These entities and the values of their properties would be linked together by declaring RDF triples.

It then becomes possible to describe the meaning, the semantic of Web pages beyond the syntax of natural languages and their inherent ambiguity. A Web of linked data enables computer agents to follow the links and perform more intelligent operations using the knowledge behind the words.

For this, the SPARQL Protocol and RDF Query Language [16] enables queries within the huge graph of RDF triples that constitutes the Web of linked data. Fig. 3 shows the state of this graph at the end of 2011 that grouped over 200 datasets, 26 billions RDF triples and 400,000 property links.



Fig. 3. The Linking Open Data (LOD) cloud diagram in 2011 [17]

Each node on the figure represents a data set. For example, the Dbpedia node at the center of the figure groups most of the information in Wikipédia, while the FOAF dataset groups information about persons having a URI on the Web. The links between two nodes means that the terms of the vocabulary in one node are linked with terms in the other node. For example, "persons" in DBpedia is related to "persons" in FOAF and their geographical localization can be found in another vocabulary such as GEONAMES.

In the same way, terms in ISO-MLR are linked to terms of other vocabularies on the Web of data.

For example, iso-mlr5:Person in the graph of Fig.2 has the same meaning as foaf:person or dcterms:person. This means

that a computer agent that would search for an isomlr:learning_resource can also retreive its isomlr:Contributors, find these persons and retreive their Wikipedia pages from Dbpedia, their email from FOAF and their localization from GEOBASE.

IV. COMETE, A RDF-BASED RESOURCE MANAGER

COMETE is a learning resource repository manager based on the RDF approach. It allows locating, aggregating and retrieving educational resources that constitute the heritage of an organization. Basically, it is a database containing metadata about learning resources on which users can perform queries to find and discover educational material that they can reuse for their various needs.

Fig.4 describes the technical architecture of a COMETE implantation instance. It's a 3-tiers client-server architecture developed in Java technology.



Fig. 4. COMETE architecture schema

Various web applications powered by an Apache Tomcat server provide specialized REST services that allow different types of clients to exploit the open data contained in the repository. Most of the clients use their favourite web browsers to access the system through a user-friendly web interface. A SPARQL endpoint is also available for advanced users who want to directly access the raw RDF triples to built various applications or Web services.

A. Integrating new resources in the triple store

The integration of resources inside a COMETE repository is done by imports of their metadata records. There are different ways to achieve this task. The metadata records can be imported manually by uploading an archive file containing a collection of metadata records. Most of the time, however, the metadata records will be harvested automatically by either an OAI-PMH Harvester or a HTML Spider. In such a case, a Harvest Definition will declare the technical information required to access the repository to be harvested. Using the facilities provided by the operating system on which COMETE runs, it's also possible to program harvest schedules so that the process is executed periodically to make sure that new or updated metadata records are always imported to the system.

These records are ingested by the system and a XSL transformation extract data for generating all pertinent triples. COMETE enable data mining across multiple metadata schemas like Dublin Core, IEEE LOM and other application profiles. Fig. 5 illustrates the result this process, that is a homogeneous graph of data in accordance with COMETE's internal metamodel (Fig. 6).



Fig. 5. Transformation from metadata to RDFgraph

All the generated triples are stored into Mulgara [18], an open source RDF triple store system, where data is organized around different RDF graphs. The default graph contains all the triples about learning resources whereas some other specialized graphs manage SKOS thesaurus and other different views of the system.



Fig. 6. COMETE metamodel (main classes)

As a semantic network, the RDF graph represents the model entities as nodes. Mains nodes are learning resources (LearningObject), persons and organizations (Identity) and element of vocabulary (SKOS Concept). By various techniques, the system tries to maximize the inner coherence of the graph.

The Identity module (on Fig. 4) implements the management of identities (metadata about persons or organizations). This includes importation of identities, identity resolution (identifying any identity in the database that represents the same person or organization, making sure it

stays unique, and completing it as new details are known), and representations of the identity (such as VCards and HTML code for the end users). Furthermore, manual merge of identities is also provided within a set of administrative tools for a better control of data integrity.

The Vocabulary module (on Fig. 4) implements the management of vocabularies and thesauri, which involve importing from VDEX or SKOS formats, unambiguously identifying the vocabulary that a term is from, and finding a computer readable representation of the whole vocabulary, updating from source automatically, transparently converting from one format to another, replacing a vocabulary when updates are available, publishing vocabularies automatically and providing user interface elements reusable by other modules, such as efficient vocabulary term choosers for queries to the repository.

This module manages also correspondences between taxonomies. Indeed, SKOS concept alignment between different ontologies (or vocabularies) can be taken into account by the query engine. A useful example of alignment is the mapping between different school-level taxonomies of different countries to promote the interoperability of resources between national repositories. For instance we can search resources which target audience is Junior High School in United States and the results may contain pertinent Secondary School I-III tutorials produced in Québec. We can imagine here a wide range of possibilities.

B. Querying the triple store

All of the previously presented modules provide rich graphs of data that allow doing more sophisticated searches in the repository. All nodes have many textual information (literal triples) where values are indexed and which can be used by fulltext search methods.

Nodes are also linked together and their graph can be traversed to perform more "intelligent" searches. Furthermore, all of the elements in the model are referenced with a unique identifier (URI).

The next figure represents the simplest COMETE search mode. As in a Google search, it only needs a field of keywords.



Fig. 7. COMETE Simple Search

Suppose now we seek the resources of an author X dealing with Organic Chemistry from Dewey classification. COMETE web interface offers different kinds of search interfaces, including advanced search. Fig 8 illustrates the easy way to fill the previous query.

| sinp | Ne Search | Auvanceu Search Inemau | <u>. Navigat</u> | | Collections |
|------|-------------|------------------------|------------------|---|-----------------|
| Sean | ch of resou | from contributor | ~ | 1 | Peter Greenaway |
| | | | | _ | |

Fig. 8. COMETE Advanced Search

Search

All of the queries will be translated in SPARQL language by the QueryEngine module of Fig.4 and then be run on the triple store. By combining different conditions, mixing keyword-based approach and by using negative prefixes, more complex queries can be performed.

A third way to navigate inside resources is to use the Thematic Navigation (Linked data on Fig.4) module that lets user discover resources from available thesaurus.

| Dewey | | | | | |
|-----------------------|---|--|--|--|--|
| Dewey • 🕨 Science • 🕨 | Chemistry Physical chemistry | | | | |
| ₽ Search | Chemistry & allied sciences Physical chemistry Techniques, equipment & materials Analytical chemistry Inorganic chemistry Organic chemistry Crystallography Mineralnov | | | | |



Results will be returned and displayed to the user interface. In connection with alignment of vocabularies seen previously, queries may be extended with option like: "include equivalent categories from Library of Congress". Obviously, resources linked with that classification must be present to obtain more results.

Finally, a fourth search mode in COMETE user interface is the Collection mode. It offers to users a list of preset complex queries to avoid having to deal with. Example: the "Last month Algebra and Euclidean geometry resources from Montréal University" set.

Like we said previously, there is also a SPARQL endpoint for querying the triple store. To achieve this task, a Snorql module is deployed. It offers a user web interface with a text area field to enter and run SPARQL queries. Although it requires some technical knowledge, it's still a simple way to explore data.

C. Linking with the Web of data

The link with the web of data, as a global data space, is done by the following facts:

COMETE respects the basic principles of Linked Data:

- ✤ All objects are described with HTTP URIs.
- URIs can be dereferenced over the HTTP protocol into a description of the identified object or concept. Moreover, the content negotiation mechanism (or 303 redirect strategy) makes possible to serve different versions depending on the context of the client; HTML page for web browser clients, RDF/XML for software agent.
- Links are included to other URIs as soon as possible to promote future discovering of resources.
- ✤ A unifying RDF data model provides a globally unique identification of entities and allows different schema to be used (and reused) in parallel to represent data.

This last bullet point is supported by a vocabulary reference [19] that details the COMETE meta-model in terms of class and property definitions and the reuse of existing vocabularies (Dublin Core, FOAF and SKOS).

The publishing of data via a SPARQL endpoint allows interaction with COMETE data by external systems.

V. USING COMETE WITHIN A MOOC PLAFORM

In this section, we present two use cases where COMETE is used to interoperate with a MOOC platform like OpenEdX. Within such a platform, the role of COMETE is twofold: 1) enable designers to search and reference OERs within a MOOC; 2) reference MOOC themselves to produce a searchable standardized MOOC portal.

A. The OpenEdX MOOC platform

OpenEdX is the open-source release of the edX platform developed by a non-profit organization founded by Harvard and MIT in the USA. In April 2013 Stanford and edX agreed to collaborate on future developments of the edX platform. In September 2013, Google committed to the development of OpenEdx. In France, France Université Numérique (FUN) uses a version of OpenEdX. Amongst many other institutions, Téléuniversité du Québec has also adopted OpenEdX for its MOOC initiative.

OpenEdX provides essentially two server-based applications. The first one, edX-STUDIO, is the application where designers build courses. Resources and activities are grouped in course modules and stored in Mongo no-sql XML files and in MySQL databases. Students interact at runtime with a second application, the Learning Management System (LMS) that performs learner authentication, runtime learning scenario support, forums and online group meetings, automatic and peer-assessed grading of learners and learning analytics operations [20].

B. Designing a MOOC using the COMETE OER Manager

Figure 10 present an example of a course structure. The course is subdivided in sections (e.g. modules) and each module in sub-sections (e.g. lessons). For each lesson, the upper bar provides access to the lessons' sequential components. It is can be composed of four kind of components: discussion components, HTML components, problem components, and video components. In principle, all these components should be open educational resources (OER).

All these OER components re found mainly on the Web. Actually. Designers use search engines like Google or Bing to find open resources to reuse or adapt for their course. As explained previously in section II-D, there are many adavantages in using a learning resource repository manager like COMETE to find suitable ressources.

Using REST web services, a call to COMETE from OpenEdX studio could start efficient search operations and facilitate the selection of ressources of the four categories proposed in STUDIO. Conversely, STUDIO could be upgraded to provide forms to edit metadata for the resources in a standardized DC, LOM or ISO-MLR application profile suitable for Studio. This would enable designers to automatically create a resource repository for a course, for a whole program or for all edX users. The creation of this local repository would produce a URI where the edX ressources can be harvested by COMETE or other OER Managers and integrated into larger repositories for future use.

| TELUQ: INF-6070 I | ntroduction au V | Veb sémantique | | | | ۵ | Gilbert 👻 |
|---|------------------|---|----------|---|--|------------|--------------------------|
| Courseware Course Info Travaux notés Équipe | Discussion | Wiki Progress | Présent | ation du cours Feuill | e de route Res | sources En | icadrement Staff view |
| Apprentissage Module 1 | | ۵ | | ۵ | | 8 | F |
| Module 2 Module 3 Module 4 Module 5 | | <mark>iule 1</mark> epts de bas aissances | E EN RE | PRÉSENTATION E | T GESTION D | ES | |
| | Aperçu | des activités | | | | | |
| | | Activités | Semaines | Ressources principal | es Productions | Notation | |
| | 1.1 | Le démarrage du cours | 1 | Présentation du cour (Vidéo) <u>Contenu du cours</u> (Vidéo) | Premier contact avec le chargé d'encadreme | ent | |

Fig. 10. A screen from OpenEdX LMS

C. Referencing a MOOC using COMETE

When a new MOOC is created in OpenEdX, a screen like the one on Fig. 11 is offered to the designer. Actually, only four metadata are asked: the course name, the organization that supports it, the course number and the periods when it will be offered.

The form shown on figure 11 could be easily extended to fields from a DC, LOM or ISO-MLR application profile that would take in account the differences between small resources within a MOOC, compared to large OERs like complete MOOCs.

| ly Courses | |
|--|--|
| elcome, Gilbert! e are all of the courses you currently have access to in Studio: | |
| Create a New Course | |
| e.g. Introduction to Computer Science | The public display name for your course. |
| Organization * | |
| e.g. MITX or IMF | The name of the organization sponsoring the course Note: No spaces or special characters are allowed This cannot be charged |
| Course Number * | This cannot be changed. |
| e.g. CS101 | The unique number that identifies your course withi your organization - Note: No spaces or special characters are allowed. This cannot be changed. |
| Course Run * | |
| | The term in which your course will run - Note: No |

Fig. 11. Creating a new MOOC in edX STUDIO

Then, automatically, each time a new MOOC is created, it would have a URI on the Web of data together with its composing resource. COMETE could them provide a searching facility is MOOC repositories like Class Central [21], which is a free online MOOC aggregator from top universities like Stanford, MIT, Harvard, etc. offered via Coursera, Udacity, edX, NovoED, & others.

Actually, in a MOOC portal, courses are classified by subject, universities, level and provider, which are the only meta-data entries available to browse for a course. Most of the time one must open each course to know what's in it. With standardized metadata, COMETE could power a MOOC portal with various kinds of search and navigation within repositories containing hundreds of MOOC, including advanced search combining many metadata and navigating on the Web of data.

CONCLUSION

We have presented a solution to one of the main problems in Open Educational Resources repositories, which is the multiplicity of norms, standards and application profiles that preclude efficient search for resources within multiple repositories. We have built a first Linked data OER repository manager, COMETE, relying on semantic web techniques, largely complying to the new ISO-MLR emcompassing and flexible standard. Also, its use for MOOC and MOOC components referencing using RDF triples will become an asset as massive online courses are growing rapidly in most countries. Our next work will be to investigate various integration of COMETE tools with MOOC platforms as indicated in the present contribution.

ACKNOWLEDGMENT

The authors wish to thank Frédéric Bergeron who has strongly contributed to the development of COMETE. Also, we must underline the implication of Marc-Antoine Parent, Benoit Grégoire, Gilles Gauthier and Pierre-Julien Guay who have provided advice on some of the orientations of the tool. We also thank Claude Coulombe and Mohammed Ben Jemia for their implication with the OpenEdX MOOC platform.

Finally we wish to aknowledge the financial support of the Ministère de l'Enseignement Supérieur de la Recherche, de la Science et de la Technologie (MESRST), Télé-université (Centre LICEF) and the Collège Bois-de-Boulogne (Vitrine Technologie-Éducation).

REFERENCES

- Paquette G, Lundgren-Cayrol K, Miara A, Guérette L (2004) The Explor@2 Learning Object Manager, in R. McGreal (ed), Online education using learning objects. pp 254-268. London: Routledge/Falmer.
- [2] Paquette, G. (2010). An ontology-driven System for e-learning and knowledge Management. In Paquette, G., Visual Knowledge Modeling for Semantic Web Technologies: Models and Ontologies. Hershey, PA: IGI Global, pp 302-324
- [3] Allemang D. and Hendler J. (2011) Semantic Web for the Working Ontologist – Effective Modeling in RDFS and OWL. 2nd Edition. Morgan-Kaufmann/Elsevier, Amsterdam.
- [4] Domingue, J., Fensel, D. et Hendler, J. A. (dir.). (2011). *Handbook of semantic web technologies*. Berlin, Allemagne : Springer-Verlag.
- [5] Heath, T. et Bizer, C. (2011). Linked data: Evolving the web into a global data space. In Synthesis Lectures on the Semantic Web: Theory and Technology, 1(1), 1-136.
- [6] ISO-MLR (2013) ISO-IED 19788 Information technoogy Learning, education and training – Metatda for learning resources multipart standard. http://en.wikipedia.org/wiki/ISO/IEC_19788.
- [7] Hollands, F. and Tirthali, D. (2014) *MOOCs: Expectations and Reality.* http://cbcse.org/wordpress/wp
 - content/uploads/2014/05/MOOCs_Expectations_and_Reality.pdf
- [8] Daniel, J. (2012) Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education*. http://www-jime.open.ac.uk/article/2012-18/html
- Unesco (2012). 2012 Paris OER Declaration. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/WPFD20 09/English_Declaration.html.
- [10] Duval, E. and Robson, R. (2001) Duval, E. and Robson. R. Guest Editorial on Metadata. *Interactive Learning Environments, Special issue: Metadata*, Volume 9-3, December 2001, pp. 201-206
- [11] DC Dublic Core Metatdata initiative. http://dublincore.org .
- [12] IEEE-LOM Learning Object Metatda, http://fr.wikipedia.org/wiki/Learning_Object_Metadata .
- [13] GLOBE Global Learning Object Brokered Exchange, http://globeinfo.org.
- [14] OAI-PMH Open Archives Initiative Protocol for Metadata Harvesting. http://en.wikipedia.org/wiki/OAI-PMH.
- [15] Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American Magazine, vol. 284, no. 5, mai 2001, pp. 29–37
- [16] SPARQL SPARQL 1.1 Query Language, W3C Recommendation, 21 March 2013. http://www.w3.org/TR/sparql11-query/
- [17]
 LOD Cloud The Linking Open ata cloud diagram. http://lod-cloud.net

 [18]
 Mulgara
 RDF
 Triple
 Store
 System.
- [18] Mulgara RDF Itiple Store System. http://en.wikipedia.org/wiki/Mulgara_(software)
- [19] Miara A. and Bergeron, F. *COMETE RDF Metamodel*. http://comete.licef.ca/reference/
- [20] Coulombe C. (2014) *Expérimentation de la plateforme OpenEdX*, rapport technique LICEF, Télé-université du Québec.
- [21] Class Central (2014) Free Online Education Portal, https://www.class-central.com