AraCore: an Arabic Learning Object Metadata for Indexing Learning Resources

Hend S. Al-Khalifa, Hugh C. Davis
Learning Technologies Research Group
Electronics and Computer Science School
Southampton University
United Kingdom
{hsak04r/hcd}@ecs.soton.ac.uk

Abstract

In the era of information systems globalization the need to have educational metadata to index and describe digital learning resources for easy searching, retrieving and reusing them quickly and efficiently is becoming an essential research topic in learning technologies discipline.

In this paper we will present a brief overview of metadata standards, protocols and application profiles. Then we will discuss the issues related to the need for an Arabic learning object metadata. Also we will call for the formation of a community of practitioners to identify guidelines for metadata implementers, creators and users in the use of metadata in e-learning content among Arabs.

Finally, we propose a sample metadata application profile called AraCore which will be based on the IEEE 1484.12.1-2002 standard. This will be our first attempt to help the Arab community to think about creating an Arabic learning object metadata application profile to be used in assessing the exchange of Arabic learning objects.

Keywords

Metadata, application profile, digital repository, learning object, Arabic, AraCore

1. Introduction

Sharing and exchanging learning materials between teachers is not a new idea. In the traditional classroom, sharing was used to save teacher's time, effort and to exchange expertise in a specific subject area. It was not until the 1990s, when educational materials began to emerge in their electronic form that the recognition of the importance of reusing learning materials for economical and time saving purposes beside transferring expertise and ideas initiated the creation of metadata standards for indexing and sharing [1].

Metadata is a record that consists of structured information about a resource; it can be also defined as information about information or data about data. Metadata is basically used to annotate learning materials (objects) to simply identify and describe their content. While looking into the history of metadata it can be found that it is not a new concept. Over 2000 years ago, metadata has been used for cataloguing and indexing information stored in libraries and now its usage has been extended to index information and to provide the semantics for the semantic web[1].

Learning materials in their electronic form are called learning objects. According to IEEE, Learning Objects (LO) can be defined as "any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning" [2]. From the previous definition a learning object can be any resource with an explicit educational application. It can be digital, for example, a simple Microsoft Word, PDF, or text, an e-book, or a Flash animation. Or it can be physical like a textbook or CD-ROM. But the concern will be on digital representation of learning materials, due to the fact that they can be easily distributed and shared using a network, while physical learning materials do not have this capability [3].

To keep track of learning objects dissemination, they need to be stored in a special database called a Learning Object Repository (LOR) or for short digital repository. The role of LOR is to aggregate a collection of learning objects for a defined community and store them in a single location. CAREO (http://careo.netera.ca) , Merlot (www.merlot.org),

[1] http://www.w3.org/2001/sw/

EdNa (www.edna.edu.au), and RDN (www.rdn.ac.uk) are examples of digital repositories with customized learning objects metadata (LOM).

2. Evolution of educational metadata from standards to application profiles

Two international metadata standards are widely used to index learning objects: the Dublin Core Metadata Element Set (DCMES) and the IEEE Learning Object Metadata (LOM) Standard. The later is more popular due to its widely acceptance in e-learning communities. In addition, the main difference between the two is that; IEEE LOM was originally developed specifically for the domain of education and training while the DCMES was originally developed for general resources, and is now being adapted for the specific field of education and training.

Correspondingly, as more and more applications are implemented using educational metadata, it becomes obvious that it would be difficult for a single metadata model to accommodate the functional requirements of all applications (Chatzinotas el al, 2004) [5]. Because of that application profiles have emerged.

Sampson (2004) [4] defines an application profile as "an assemblage of metadata elements selected from one or more metadata schemas." Those, the application profile will serve as an adaptor of a particular metadata schema or multiple schemas and it will be tailored to the functional requirements of a particular application taking into account interoperability[2] with the original base schemas [4, 5].
Typical examples of well-known application profiles include CanCore [6], Ariadne [7] and UK LOM [8].

[2] Interoperability can be defined as the ability of software and hardware on different machines from different vendors to share and use data (from www.webopedia.com).

3. Protocols for exchanging data between digital repositories

In an interesting article by David Haynes (2004) about the five purposes of metadata, he mentioned that one of the purposes for using metadata is to allow assessment of compatibility and exchange of information between systems [9].

Therefore, metadata plays a key component in interoperability; and to achieve that IMS[3] has initiated a specification called Digital Repository Interoperability (DRI) to define a set of protocols for interoperability between digital repositories.

One of these protocols is XQuery[4] over SOAP[5], which uses XQuery transactions to search and retrieve learning objects from digital repositories. Z39.50 protocol is another protocol recommended by IMS DRI. It is used for federated search across several libraries. The third recommended protocol is OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) [10]. This protocol allows service providers to harvest metadata from data providers. The main difference between the last two protocols is that Z39.50 is a cross-system search protocol while OAI-PMH is a metadata harvesting protocol.

Moreover, all these protocols use metadata as the bases for searching/harvesting/retrieving learning objects from digital repositories. For more information about IMS DRI see [11].

Finally, it is worth mentioning that there are common communication frameworks for querying digital repositories. For instance, Simple Query Interface (SQI) [12] based on the work by Ariadne, CELEBRATE, Edutella, Elena, EduSource, ProLearn, Universal/EducaNext and Zing, and eduSource Communication Language (ECL) used by eduSource are examples of application programming interfaces (API) for querying federated digital repositories. These frameworks rely on metadata as a transport vehicle for their quires.

[3] IMS Global Learning Consortium is an organizational body that supports the adoption and use of learning technology worldwide by promoting the adoption of open technical specifications for interoperable learning technology. (www.imsglobal.org)

[4] http://www.w3.org/TR/xquery/

[5] http://www.w3.org/TR/soap/

4. Why an Arabic application profile

In our previous overview we draw some insight about the significant places where metadata plays an important role (e.g. tagging learning materials, indexing LOs in digital repositories and building up queries for searching/harvesting protocols). It seems quite obvious and convincing that for Arabs to be fully involved with learning technology revolution; a tailored version (i.e. an application profile) of IEEE LOM is needed to fulfill the functional requirements of the Arabic community and their language needs. In addition, the Arabic application profile is needed to bridge the gap between the Arab countries and others in the field of learning technologies.

Therefore, to answer the question "*why do we need an Arabic application profile*?" three important reasons need to be identified: the Arabic language in computers and the web, Arabs educational system and finally the existence of Arabic learning objects on the web.

4.1 Arabic language in computers and the web

The Arabic language is the mother tongue of over 300 million people in 22 Arab countries. The majority of Arabs, particularly in Saudi Arabia, Egypt, United Arab Emirates, Kuwait, Bahrain, Qatar, Oman, and Syria use Arabic as the first language in their educational system.

The nature of the Arabic text, its direction in writing (i.e. Arabic text is written right-to-left) and the shapes of its characters depend on their position in the word. The Arabic script change the shape of the letter depending on its position in the word and what other letters precede of follow it in the word [13]. These aspects make displaying and printing of Arabic more complicated and often challenging than other Western languages.

Moreover, there is the problem of Arabic encoding on the internet. The Arabic language on the internet has gone through a lot of changes when the internet first introduced in the Arab region. At the beginning the Arabic text was displayed using a GIF or JPG file. This was because web browsers can not display Arabic properly. Then as Microsoft began dominating the browsers market it produced the so called (Windows code page 1256) encoding for the Arabic language. Yet, not all Arabic web sites use Windows CP-1256;

there have been verities of them that use Unicode (UTF-8) encoding. For further information about the history of Arabic text on the Internet please visit [14].

4.2 Arabs educational system

Historically and culturally, the Arab region has many features in common, in particular from the religious standpoints. On the other hand, there is a slight difference in curriculum and in the level of education. For instance, in Saudi Arabia they teach four main religious subjects and the textbooks used which primarily cover the traditional religious texts and their interpretation, change very little over the years. While, textbook materials in fields such as mathematics, science, and social studies are reevaluated regularly. While most Arab countries have only one religious subject and the textbook in use is subject to modification or change.

Furthermore, similar textbooks are used by male and female students who also follow the same academic curricula. Thus many Arab countries have schools devoted for girls as well as for boys in the primary and secondary education level, while at the level of higher education most Arab counties have a mixed environment.

4.3 Arabic learning objects on the web

Additionally, there is some existence of Arabic learning objects on the internet created by some members of the community, which in turn were used as informal learning resources by their peers. One example is the site designed by the ministry of education, center of educational technologies in Saudi Arabia (http://www.moe-edc.org/a/tec/etc/link4.asp). The site consists of number of flash movies explaining the complete science textbook curriculum for the fourth grade.

In addition, there are many Arabic educational forums like (http://www.moudir.com/) which promotes the exchange of learning materials such as PowerPoint slides and exam templates in word documents. That means that the Arab community have their own learning objects scattered throughout the web in an arbitrarily fashion and it requires some sort of an organization to store them in digital repositories so that teachers and learners can reach learning materials very quickly and consistently.

4.4 The need for an Arabic application profile

From the previous discussion it seems that the need for an Arabic application profile needs to be considered seriously. Arabs can not simply adopt or translate an existing application profile like CanCore or UK LOM because it will not suit Arab requirements, due to the difference in education system structure, culture and curricula. To give an example, in UK LOM the element 5.6 named context has the following value space (*nursery education, primary education, secondary education, sixth form college, further education, higher education, continuous professional development, vocational training, community education*) while for instance in the Saudi educational system the value space will be (*primary education that consist of "primary and intermediate School", secondary education that consists of "general secondary school, religious secondary school and technical secondary school", higher education, further technical and vocational Training, teacher training*) [15].

Similarly, the LOM (IEEE 1484.12.1-2002) is culturally biased based on a claim by Blandin (2004) [16]. He stated that IEEE LOM does not cover all the learning situations, for example the value space for the element 5.1 (Interactivity Type) is {active, expositive, mixed}, while other types of interactivity have been missed like "interrogative" or "behaviourist". Also, the element 5.5 (Intended End User Role) has the value space {teacher, learner, author, manager}. These values correspond to the roles in the "Instructionism" paradigm. Tutor, mentor or coach, do not appear in the list. This implies that IEEE LOM is designed for a specific learning paradigm and it did not take into consideration other learning paradigms like behaviorism for instance [16].

5. Steps proposed for creating an Arabic application profile

To propose an Arabic application profile a considerable debate is required in order to identify what kind of metadata is actually needed and how granular any future of the metadata specifications should be.

By searching the internet, the authors have been unable to find any sign of an Arabic community or a national body interested in applying e-learning standards in the Arab world. For this reason and to start the process of building the AraCore LOM the

following steps are proposed. Note that these steps have been inspired by the UK LOM draft [8]:

1- Call for the formation of a community of practitioners to identify common Arab practice in the use of metadata in packaged E-learning content.

2- Comparison of metadata schemas based on the IEEE LOM standard (e.g. CanCore, UKLON and SigCore) to see which one can become the ideal schema to start with.

3- After comparison is done a set of guidelines must be drafted on the implementation of a minimum common core of LOM elements and associated value space. Also call for communities to create controlled vocabulary available online.

6. Difficulties and challenges encountered in creating an Arabic application profile

Many difficulties and challenges will face the development of an Arabic metadata application profile; which include:

1- No communities exist yet either locally (country-wide) or nationally (nation-wide).

2- No digital repositories dedicated for archiving learning resources exists yet.

3- The need for governmental support (ministries of education, higher education and universities (public and private)).

4- Building applications that utilize the creation of Arabic LOM.

5- Having a reliable internet/network infrastructure between schools and educational institutes is not yet implemented.

6- Finally, the lack of evangelists in the Arab world and most important is realizing the importance of metadata to enhance learning.

7. AraCore: the Arabic application profile

To propose a starting point for the Arabic Learning Object Metadata (LOM) application profile, the underpinning standard and its characteristics need to be explained.

Similarly, the definition of the Arabic LOM application profile (which we will call "**AraCore**") and its purpose should be stated clearly as follows:

> **AraCore**: is a metadata application profile derived from IEEE LOM international standard tailored to fit the needs of the Arab educational system and its language and to support learning resources management, description of educational purpose, interoperability and accessibility.

According to the previous definition AraCore will be based on the IEEE LOM standard. The IEEE LOM has 9 categories each of which is optional. Each category relates data elements that cover a specific aspect together; for example the technical category covers things like the file type, date of creation and so forth. One of the main features of IEEE LOM is its ability to extend and add new data elements upon application needs. This flexibility in the standard encouraged LOM developers to use IEEE as the base standard for developing new application profiles that suits their application needs. Also the value space for some of the elements will be derived from either IEEE value space or will be self-developed vocabulary.

As a first attempt to create the AraCore application profile, we have translated the original IEEE LOM standard into Arabic and proposed some value spaces and guidelines so the Arab community can get started. Notice that it is just a seed for the AraCore application profile till we have formed a community. A first draft of AraCore application profile can be downloaded from [17].

8. Conclusion

The AraCore metadata initiative from our point of view will play an important role in bridging the gap between Arab world and the developed world in terms of learning technologies utilization and opens to other expertise and experiments in educational field.

The success of such an initiative will benefit the Arab world in the following ways:

- Sharing and reusing of learning materials between Arab countries;

- Exchanging non-textual learning objects such as images, sound tracks, and executable Java applets with other international LOR;
- Economically feasible in the long run in terms of time and efforts; and
- Encouraging members of the community to create digital resources to utilize the usage of computers and the internet.

References

1. McClelland, M., *Metadata standards for educational resources.* 2003. **36**(11): p. 107.
2. 1484.12.1: IEEE Standard for Learning Object Metadata. Accessed on 2 May. 2005.Available on line http://ltsc.ieee.org/wg12/
3. Stratakis, M., V. Christophides, K. Keenoy, and A. Magkanaraki, *E-Learning Standards*. 2003. p. 45.Available on line http://www.dcs.bbk.ac.uk/~ap/projects/selene/reports/seleneELearningStandards.pdf
4. Sampson, D. *The evolution of educational metadata: From standards to application profiles*. 2004.
5. Chatzinotas, S. and D. Sampson. *EMAP: Design and implementation of educational metadata application profiles*. 2004.
6. CanCore. Accessed on 25 May 2005.Available on line http://www.cancore.ca/
7. Ariadne. Accessed on 22 May 2005.Available on line http://www.ariadne-eu.org/
8. UK LOM. Accessed on 25 May 2005.Available on line http://www.cetis.ac.uk/profiles/uklomcore
9. Haynes, D., *The Five Purposes of Metadata*, in *Update magazine*. 2004, Chartered Institute of Library and Information Professionals (CILIP).Available on line http://www.cilip.org.uk/
10. Open Archives Initiative Protocol for Metadata Harvesting. Accessed on on line http://openarchives.org
11. IMS Digital Repositories Interoperability - Core Functions Information Model. Accessed on on line http://www.imsglobal.org/digitalrepositories/driv1p0/imsdri_infov1p0.html
12. Simon, B., D. Massart, F.v. Assche, S. Ternier, E. Duval, S. Brantner, D. Olmedilla, and Z. Miklos. *A Simple Query Interface for Interoperable Learning Repositories*. in *Proceedings of the 1st Workshop on Interoperability of Web-based Educational Systems*. 2005. Chiba, Japan.
13. Diab, H. *Standardization Related To Arabic Language Use in ICT*. in *The Western Asia Preparatory Conference for WSIS*. 2003. Beirut, Lebanon.
14. Arabic on the Internet. Dynasty, B. Accessed on 15 July 2005.Available on line http://baheyeldin.com/
15. Education in Saudi Arabia. Accessed on 24 May 2005.Available on line http://www.wes.org/ewenr/01nov/practical.htm
16. Blandin, B. *Are e-learning standards neutral?* in *International Conference on Computer Aided Learning in Engineering Education*. 2004. Grenoble, France: CALIE 04.
17. AraCore specification. Al-Khalifa, H. Accessed on 30 May 2005.Available on line http://www.ecs.soton.ac.uk/~hsak04r/