

Protecting the content through learning object metadata

Shaina Raza, Syed Raza Bashir¹

Department of Computer Science, Ryerson University, Toronto, Canada, ¹Department of Computer Science and Engineering, University of Engineering and Technology (UET), Lahore, Pakistan

Abstract

The web is full of numerous educational resources but they are not being properly used by the educators. There is so much pedagogical content available on the open web that is being ignored. A lot of learning initiatives stepped in to propose recommendations and guidelines to ensure interoperability of digital content. This has led to the development of learning objects repository (LOR) whose goals are interoperability, reuse, sharing, and retrieval of learning content. However, at the same time, the reproduction of learning material should not breach the copyright protection of the right holders as it is an act of cybercrime. In the lifecycle of LOR development, learning objects (LOs) are annotated using metadata descriptors to specify their syntax and semantics. This annotation process has led to the development of learning objects metadata (LOM) whose ultimate goals are to make searching and cataloging of LOs an easier task. LOM standard includes a number of sections, one of which is the “Rights” category which takes care of intellectual property rights and terms for the use of an LO. This paper presents the idea that how learning resources are annotated using LOM standard and how this annotation contributes to anti-cybercrime in formal education. More specifically, the paper tells that the “Rights” category and some related elements that work together for the provision of protection to the content holders. The paper also suggests that there should be some standardized mechanism for the automatic annotation of LOs so as to give copyright protection on permanent basis.

Key Words: Annotation, e-learning, learning management system, repository

Address for correspondence:

Dr. Shaina Raza, Faculty of Information Technology, Ontario College for Development Training, Toronto, Canada. E-mail: shainaqamar@hotmail.com

Received: 11.04.2016, **Accepted:** 26.04.2016

INTRODUCTION

The web is full of numerous educational resources but they are not being properly used by the educators. Hence, many pedagogical resources online are ignored because they are not being accessible, interoperable, and reusable in the real sense. Before the e-learning community invests so much cost, time, and energy into building qualitative e-learning content, it must consider that the open web pedagogical resources can be easily loaded into the learning content management or web portal for

the delivery and access of learning content. As per the Institute of Electrical and Electronics Engineers’ (IEEE) definition, each such content item is a learning object (LO) and can be defined as “any entity, digital, or non-digital that may be used for learning, education, and training.”^[1]

However, the reproduction of the learning content must seek permissions from the right holders. Citation is a method of validating the original source so as to endorse the honesty to

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Raza S, Bashir SR. Protecting the content through learning object metadata. *Imam J Appl Sci* 2016;1:3-8.

Access this article online

Quick Response Code:



Website:

www.e-ijas.org

avoid plagiarism. LOs economy leads to the reusability of the learning content; however, at the same time, the subsequent authors should not forget the republication ethics. They must respect the contribution of the right holders who had played a role in the development of pedagogical resources. The violation of the copyright protection is an act of cybercrime as it is the responsibility of the subsequent authors to pay contribution to the original source. If the copyright protection is not given to the authors or content creators, then most probably they will refrain from producing new material.

E-learning has emerged as a business in the field of education, more specifically in higher education. A lot of institutional, governmental, and industrial boards stepped in to propose recommendations and guidelines to ensure interoperability of digital content. This has led to the development of LOs repository (LOR). LOR is a digital repository that enables access to LOs at local or distributed level.^[2] The LO defined as any learning entity for the purpose of education should be openly available, reusable, and interoperable. However, at the same time, the content developers must pay tribute to the original source for each and every chunk of the content being used. Besides being a topic of software ethics,^[3] the principle of referring the original source should be a mandatory activity during the development of LOs. The reusability of the content is the primary purpose of LOs development, but the content developers should not sacrifice protection of the right holders at the cost of LOs economy.

As far as LOR is concerned, the content producers do not forget to refer to original sources as it is a part of metadata annotation. LOs are annotated to specify their syntax and semantics. This makes searching and cataloging of LOs an easier task. To accomplish the previously mentioned operations, for example, searching, reuse, interoperability, and sharing on LOs, it is necessary to annotate them with descriptive tags. These datasets are named as metadata and are standardized. The most commonly used specifications of metadata are LOs metadata (LOM) standard from the IEEE-Learning Technology Standardization Committee.^[2]

The conceptual data schema of LOM consists of the following nine sections arranged in hierarchical tree structure.^[1,2]

These categories are as shown in Table I.

As per the agenda of the conference, the category that is of primary interest is the “Rights” section.^[3] It preserves the copyright issues and other restrictions. This section of LOM standard describes the permissions for access, use, reuse, creation, publishing, and editing of learning resources.

Table 1: IEEE-learning object metadata elements

IEEE-LOM elements
General
Lifecycle
Meta-metadata
Technical
Educational
Rights
Relation
Annotation
Classification

LOM: Learning object metadata, IEEE: Institute of Electrical and Electronics Engineers

Therefore, this category directly or indirectly contributes to the anti-cybercrime knowledge. The paper shows that how can an LO be given copyright protection using LOM standard. Furthermore, it has been suggested that this activity of annotating an LO should be made a compulsory activity so as to avoid any act of cybercrime in formal education.

The paper consists of following sections. The first section gives the introduction and the second one discusses the related work. The third section discusses about anti-cybercrime in education. The fourth section is about the LOM standard and the fifth one shows the annotation of an LO as an example. The sixth section describes the “Right” category of LOM and explains its significance for copyright protection. The seventh and eighth sections briefly describe the challenges and future expectations for the standardization process in the area of anti-cybercrime.

RELATED WORK

The issue of copyright and terms in the context of LOs has been discussed in the literature. The study shows that the protection of LOs is a naive area where there is scope for improvement and research. Privacy is pointed to be a primary challenge in the context-aware recommender systems.^[4] The authors discuss the significance of digital content availability and reusability along with the principle need of copyright protection that must be given to the authors. However, the paper does not experiment practically to provide protection to the learning content. The idea of interconnected LORs in a peer to peer network is discussed in the LOP2P architecture.^[5] Every peer in LOP2P is an institute that contributes LOs in the network. The major contribution of the paper is the provision of a free license of use that suffices the purpose of freeing the user from the problem of misuse of content in unlawful manner.

Elliot and Sweeney reported in their paper that the permission policy for the reusability of learning content is far time-consuming procedure as compared to its exploration and adaptation.^[6] A recent study by Sinclair *et al.* surveys a number

of copyright licenses and protection rules in the area of technology-enhanced learning.^[7] One of their findings indicates the significance of open use policy for the use of LOs to provide copyright protection to the authors. Jana proposes the architecture Darkly that talks about privacy enforcement rules to the context-aware applications.^[8] The proposed framework gives privacy mechanism in multiple ways, for example, access control, algorithmic approaches to privacy control, and user inspection.

Thus, the summative efforts of all authors in this area are to promote reusability and sharing of pedagogical content along with the primary motive that how to provide the copyright protection to the principal stakeholders.

ANTI-CYBERCRIME IN EDUCATION

Cybercrime refers to a crime that evolves as a result of a computer system or network where it is used for the commitment of crime or it could be the target. Therefore, the crimes that are intended to harm individual(s) using information and communication technology are included in the subject of cybercrime. Such crimes include hacking, copyright violation, child pornography, and cracking information.^[9] The computer crimes comprise following major activities:

- Computer viruses
- Malicious code
- Identity threat
- Cyber terrorism
- Denial of service attack
- Espionage
- Phishing
- Spamming
- Spoofing.

In the scenario of e-learning, the sharing and content reuse is quite common and encouraged.^[10] However, the practice of reusability must not be involved with the violation of copyright protection which eventually makes it a cybercrime. No author would like his/her publication or learn assets to be owned by someone else without his/her consent or the provision of proper incentive. When another content developer cites the source in his/her work, then he/she is saved from being accused of plagiarism. The content writer needs to support his/her claims and for that he/she must respect the right holders. Anti-cybercrime in education emphasizes on digital rights in which the digital content is recognized as per the laws of several countries.^[3] The e-learning community tries to protect even that content that does not meet the requirements in the Copyright Act.

Now, the question is that what is to be protected? The answer to this is that any content either digitalized or nondigitalized must be protected. This includes:

- Printed or electronic books
- Images, drawings
- Audio or video recordings
- Web content
- Research papers.

The provision of copyright protection to the above-mentioned content is another vital issue that needs to be discussed. The basic rule of copyright is to seek permission from the right holder before the use of digitalizing any content such as website content, posting to a website, forwarding or using any content from newsletter, discussions boards, or blogs.^[3]

LEARNING OBJECT METADATA

The literal meaning of “metadata” is “data about other data.” In library science, the library catalog contains a set of records with elements that describe a book or other library items such as the author, title, date of creation or publication, subject coverage, and the index number. In the world of Internet, the metadata describe information about web resources. LOM is basically a data model that is used to describe a learning resource.^[11]

A number of organizations are contributing on research in the area of e-learning paradigm. The primary purpose of these initiatives is to produce guidelines and specifications to ensure these goals: Sharing, interoperability, reusability, durability, and search.^[12] A brief summary of these organizations who contribute to the development and growth of metadata specifications is given in Table 2.

All of these specifications consist of more or less same elements. A general outline of the elements for LOM standard is given in Table I. The category which particularly contributes

Table 2: Learning object metadata standards

Organization	References
IEEE-LTSC	http://ltsc.ieee.org/
CEN/ISSS WS-LT	http://www.cenorm.be/
ADL	http://www.adlnet.org/
IMS	http://www.imspj.org/c
DCMI	http://dublincore.org/
ISO/IEC JTC 1 SC36	http://www.iso.org/iso/iso_technical_committee?commid=45392
AICC	http://www.aicc.org/
CanCore metadata initiative	http://cancore.athabasca.ca/en/
IEEE-LOM standard	http://standards.ieee.org/findstds/standard/1484.12.1-2002.html
IMS application profile guidelines	http://www.imsglobal.org/ap/
UK LOM core	http://www.ukoln.ac.uk/metadata/education/uklomcore/

LOM: Learning object metadata, IEEE: Institute of Electrical and Electronics Engineers, LTSC: Learning Technology Standards Committee

in the copyright protection of the digital content is “Rights” category which is discussed in the subsequent section.

METADATA ANNOTATION

Metadata of the LOs can be generated in three ways i.e. automatic, semi-automatic and manual. In case of manual metadata annotation, the metadata is generated by the LOR maintainer. The results are quite qualitative but it's a long and tedious task. Researchers have carried much research on the automatic generation of metadata for LOs. It is basically a machine process of metadata extraction and harvesting. It uses automatic indexing techniques to add metadata descriptors to the LOs.^[13] A lot of contributors, educators and experts need to work together to develop automatic metadata generation applications. A number of techniques are proposed in the literature that suggests ways of automatic generation of metadata. These include support vector machines, fuzzy association, ontology based techniques, machine learning algorithms, neural networks, various clustering and classification algorithms.^[14] There are also some proposals for semi-autonomous annotation of learning resources in the literature. The paper, however presents the example of few learning objects that has been annotated in manual fashion.

EXPERIMENT ON ANNOTATION OF LEARNING OBJECTS

This experiment consists of the following softwares and standard:

- Apache Nutch crawler^[15]
- LomPad^[16]
- IEEE-LOM data model standard.^[2]

The system makes use of Nutch crawler to download LOs from the open web. The web crawler will find, parse, and download web pages along with its metadata and content. This will create an LOR that stores large amount of content. A massive number of content is available, so the process of annotating them begins now. Therefore, a third-party metatagging tool LomPad is used for annotation purpose. LomPad supports IEEE-LOM, CanCore, SCORM, and NORMETIC specifications. In this experiment, IEEE-LOM is used to annotate the LOs.

LomPad is connected to the repository by browsing the folder, and a connection is established between the LOs and the annotation interface. To manually annotate LOs, the draft standard for LOM^[11] is studied to understand the nine elements and their attributes are shown in Table 1. All the metadata records are encoded in XML format and RDF to define bindings of the LOM data model.^[14]

The annotation of an LO using the tool is shown in Figure 1.

The tool enables the saving of LOM in the XML format. This is shown in Figure 2.

Each LO is saved along with its metadata in a separate repository, a.k.a. LOM repository. This annotation will actually be used to assign different attributes to each learning entity. Annotating via a tool along with gold standard is very qualitative approach for adding metadata to the LOs. In the subsequent section, it is shown that how to work on “Rights” category for the purpose of protecting copyright. It is also

The screenshot shows the LomPad application window. The title bar reads "LomPad". The menu bar includes "File", "Language", "Profiles", and "Help". Below the menu bar, it says "Profile: IEEE LOM". There are nine tabs labeled "1 - General", "2 - Lif", "3 - Met", "4 - Tec", "5 - Edu", "6 - Rig", "7 - Rel", "8 - Ann", and "9 - Cla". The "1 - General" tab is active. The form contains the following fields:

- 1.1 - Identifier:** A text box with "Catalog" set to "CARS" and an empty "Entry" field.
- 1.2 - Title:** A text box containing "Context-Aware Recommendation System".
- 1.3 - Language:** A dropdown menu set to "en".
- 1.4 - Description:** A large empty text area.
- 1.5 - Keyword:** An empty text box.
- 1.6 - Coverage:** An empty text box.
- 1.7 - Structure:** A dropdown menu.
- 1.8 - Aggregation level:** A dropdown menu.

Figure 1: Annotation of a learning object using LomPad

discussed that there are other categories such as “Role” in LOM that supports the copyright protection.

The “Rights” category

The “Rights” category is particularly an interesting element in this regard. The copyright protection is that much a vital activity in digital rights^[3] that the “Rights” category is made a mandatory element in all LOM standards. This element is a part of schema in all LOM specifications such as IEEE-LOM,

CanCore, and SCORM. The set of attributes for this category and their explanation are given in Table 3.

While annotating LOs, the annotator will go through whole set of elements and will fill this important information. In case the process of annotation is automatic, then this element will be filled up by intelligent algorithmic approach. A snapshot of manual provision of “Rights” category is shown in Figure 3.

There are also other metadata elements that contribute to protect the authors’ rights in some way. For example, the “Lifecycle” element refers to the “Role” of the entity who has contributed to the provision of this content. The lifecycle element states the previous and current state of an LO and all the factors that directly or indirectly affect the state of an LO. The “Role” attribute states the stakeholders who contributed to the state of an LO. The possible specified values for this attribute^[3] are given in the LOM schema which is shown in Table 4.

One can also find attributes such as “Intended End User Role” and elements such as “Meta-Metadata” and “Relation” that support the ownership of the LO and related resources.

```
<?xml version="1.0" encoding="UTF-8"?>
- <lom xsi:schemaLocation="http://ltsc.ieee.org/xsd/LOM
http://ltsc.ieee.org/xsd/lomv1.0/lom.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://ltsc.ieee.org/xsd/LOM">
- <general>
- <identifier>
<catalog>CARS</catalog>
<entry>1</entry>
</identifier>
- <title>
```

Figure 2: XML binding of learning object metadata

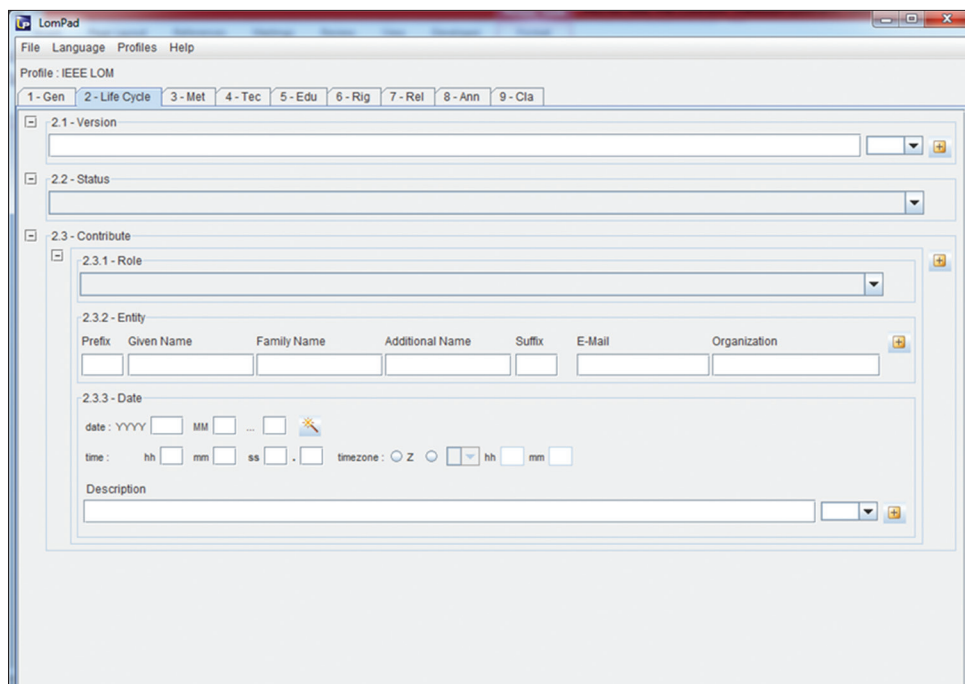


Figure 3: Annotating with “Rights” category

Table 3: Rights category and its attributes

Element	Attribute	Explanation
Rights (element #6)	Main attribute	According to IEEE 1484.12.1-2002 standards draft, the rights category is defined as: “This category describes the intellectual property rights and conditions of use for this learning object” ^[16]
	6.1 Cost	This attribute states if the use of this LO requires payment
	6.2 Copyright and other restrictions	It requires if the LO needs any copyright or other such restriction

LO: Learning object, IEEE: Institute of Electrical and Electronics Engineers

Table 4: Specified values for role attribute

Values
Author
Publisher
Unknown
Initiator
Terminator
Validator
Editor
Graphical designer
Technical implementer
Content provider
Technical validator
Educational validator
Scriptwriter
Instructional designer

CONCLUSION

In this paper, the LOM annotation of LOs is discussed. The research shows that the annotation of LOs is quite an important activity for the sake protecting the right holders. There are some elements in LOM standard, particularly with references to IEEE-LOM that contributes to the protection of digital rights. The "Rights" category particularly interests the investors and researchers in the field of anti-cybercrime. There are other elements such as "Role," "Intended End User Role," and elements such as "Meta-Metadata" and "Relation" that play important role in copyright protection. However, it can only be a fruitful activity in e-learning if the process of annotation is made a mandatory activity for the content providers.

CHALLENGES AND FUTURE DIRECTIONS

The foremost challenge for the e-learning researchers is to discover a mechanism for the annotation of LOs. Annotating LOs is nowadays a hot topic of research, and there is a lot of room in this area. The current status of LOs annotation is still under research, and no proper tools and techniques are being standardized for the annotation of LOs so far.

The first challenge in this regard is to convince educators or contributors to annotate LOs properly. The second challenge is to make it an obligatory act so as to give protection to the

right holders. Since LOM elements emphasize on authorship and related roles in a very clear manner, therefore it can significantly contribute to the anti-cybercrime in formal education.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. IEEE Learning Technology Standardization Committee, Draft Standard for Learning Object Metadata; 2001.
2. IEEE, Learning Technology Standards Committee (LTSC). "Learning Object Metadata (LOM). 1484.12.1". 2002, IEEE.
3. Rosenblatt W, Mooney S, Trippe W. Digital rights management: Business and Technology. John Wiley & Sons, Inc; 2001.
4. Verbert K, Manouselis N, Ochoa X, Wolpers M, Drachsler H, Bosnic I, Duval E. Context-aware recommender systems for learning: A survey and future challenges. *IEEE Trans Learn Technol* 2012;5:318-35.
5. De Santiago R, Raabe A. Architecture for Learning Objects Sharing among Learning Institutions—LOP2P. *IEEE Trans Learn Technol* 2010;3:91-5.
6. Elliot K, Sweeney K. "Quantifying the reuse of learning objects", *Australasian J. Educational Technol* 2008;24:137-42.
7. Jane S, Mike M, Jane Y, Stephen H. A practice-oriented review of learning objects. *IEEE Transactions on Learning Technologies* 2013;6:77-92.
8. Jana S, Narayanan A, Shmatikov V. A Scanner Darkly: Protecting user privacy from perceptual applications. In: *Security and Privacy (SP), IEEE Symposium on*; 2013. p. 349-63.
9. Moore, R. "Cyber crime: Investigating High-Technology Computer Crime," Cleveland, Mississippi: Anderson Publishing, 2005.
10. Pickard A. *Research methods in information*. Facet publishing; 2012.
11. Neven F, Duval E. Reusable learning objects: A survey of LOM-based repositories. In: *Proceedings of the tenth ACM international conference on Multimedia*. ACM. Chicago; 2002. p. 291-4.
12. Louis MK, Kindell C, Gong Q. "Automated method and apparatus to package digital content for electronic distribution using the identity of the source content." U.S. Patent No. 6,345,256. 5 Feb. 2002.
13. Duval E, Hodgins W. Metadata matters. *Proceedings of International Conference on Metadata and Dublin Core Specification, DC-2004*, Shanghai, China; 2004.
14. Vlachidis A, Binding C, May K, Tudhope D. Automatic metadata generation in an archaeological digital library: Semantic annotation of grey literature. In: *Computational Linguistics*, Springer Berlin Heidelberg; 2013. p. 187-202.
15. "Highly Extensible, Highly Scalable Web Crawler." Apache Nutch™. N.p., n.d. IEEE Standards Association. 2014. Available from: <http://standards.ieee.org/>. [Last Retrieved on 2016 Aug 25].
16. Lompad,. "Lompad". SourceForge. 2015.