Building and Nurturing Collaborative Communities

Nagesh Karmali
Email: nags@cse.iitb.ac.in
Department of CSE
IIT Bombay

July 13, 2016

Introduction

This short document highlights the paraphernalia for building and nurturing collaborative communities for creating, editing, and publishing OERs in the form of audios, videos, images, texts, manuscripts, books, and such other digital artifacts. It also identifies various existing techniques for nurturing communities.

1 Open Educational Resources Repository

Wikipedia defines Open Educational Resources(OER) as freely accessible, openly licensed documents and media that are useful for teaching, learning, and assessing as well as for research purposes [1]. OER could be in the form of any file format (images, videos, text documents, rich text documents, audio files, etc.), and as of now there is no universally accepted standard(s) for OER creation. The author creating an OER should specify an appropriate copyleft license. Some of the well known open content licenses available are Creative Commons Attribution (CC BY)[2], Creative Commons ShareAlike (CC BY-SA)[3], Public Domain Dedication (CC0)[4], GNU Free Documentation License (GFDL) [5], Free Art License (FAL)[6], etc..

2 Metadata for OER in Digital Libraries

Digital libraries have metadata for easy retrieval, and management of information resource. Metadata can describe resources at any level of composition and aggregation. It can describe a collection, a single resource, or a component of a larger resource (for example, a chapter in a book). The National Information Standards Organization (NISO) develops, maintains, and publishes technical standards related to metadata for library applications. Standards for metadata [7] in digital libraries include Dublin Core, METS, MODS, DDI, DOI, URN, PREMIS schema, EML, and OAI-PMH.

3 Building and Nurturing Collaborative Communities

Apart from static metadata dictated by the standards, the system needs additional mechanisms to dynamically tag and evolve in an ever growing fashion. To facilitate thousands of collaborative communities, a system based on crowd-sourcing techniques is essential, as manual intervention in such a system is impossible. Below, we highlight various existing components of crowd-sourcing that enables automation.

3.1 Tagging

Tagging is a way of associating a keyword to an article written on a particular subject. This allows an easy access to the article based on the search keyword provided by the user while searching for an article. It is an approach to classify articles using a bottom-up fashion as compared to the static hierarchical classification. Nevertheless, both hierarchical classification and tagging mechanism could be used in the system.
References required...

3.2 Star Rating

Star rating which is usually on the scale of 0 to 5 can be rated by the user depending on the quality of the article.
To be studied...

3.3 Voting

Voting for any article/document could be done by the user in the form of up-votes or down-votes. The number of votes defines the health of the article. Voting is used extensively in well known collaborative problem solving communities like stackoverflow[8], and many social networking sites like Twitter[9], Facebook[10], etc..

3.4 User Reputation

User reputation is based on the number of articles contributed, number of edits done by the user, appropriate reviews given for other articles, votes received for the article written, star rating received, etc.. Additional, statistics could be gathered from content versioning.

Two approaches can be employed for defining reputation scores, one is using user-driven reputation schemes, and the other using content-driven reputation schemes. Sites like Stackoverflow [8] follow the user-driven scheme, where the users are rewarded more reputation for giving good answers than for asking good questions. In [11], Adler and Alfaro show how content-driven reputation scheme could work for wikipedia and proposes a method based on the revision history of an article to compute reputation score of the contributor. In their system, authors are evaluated on the basis of how their contributions fare.

3.5 Collaborative Filtering

Collaborative filtering has a potential to harness the recommendations and behavior of users in order to make recommendations to others. Filtering helps in narrowing down the searches based on user preferences and activity. Recommendation systems are a useful alternative to traditional searches as they help users discover items they might have not have found by themselves.

4 Event Logging

Event logging is a mechanism by which all the events associated with the user is logged in the system at the server side. These logs can be analyzed for identifying the behavior/performance of the user. The origin of events can come from actions such as logging in/out, navigating, tagging, voting, reporting, etc.. The logging of events allow data analytics to be performed on millions of users using the system. Such functionality has been seen in recent MOOC platforms like open edX [12].

5 Content Versioning

Content Versioning is a kind of system that tracks and provides control over changes to the content. It provides the flexibility to the contributors for working collaboratively on editing the document. Various tools available for versioning are github [13], SVN (Subversion) [14], CVS (Concurrent Version System) [15], etc.. Content versioning available in some systems like Wikipedia, is powered by Mediawiki[16].

6 Plagiarism Detection

Plagiarism detection tools for content checking could be built in the system for identifying plagiarized content submitted by the user. This can notify the users for their inappropriate behavior. Plagiarized content submitted by the users can reduce the reputation score of the user.

7 Power of Event Logging and Content Versioning

Enhanced data analytics could be achieved using event logging and content logging mechanisms together in the system. An event cannot capture the amount of content contributed by the user, which can be easily handled by content versioning mechanism. This information can be used to rank users based on their contribution towards the subject. Users who keep a track of content being modified can best be analyzed by the event logs. Thus, both logging mechanism should be an integral part of the system in order to build and nurture growing communities.

References

[1] "Open Educational Resources: A Wikipedia Article. Available at "https://en.wikipedia.org/wiki/Open_educational_resources". Accessed on July 13, 2016."

[2] "CC BY License. Available at "https://creativecommons.org/licenses/by/3.0/legalcode". Accessed on July 13, 2016."

[3] "CC BY-SA License. Available at "https://creativecommons.org/licenses/by-sa/3.0/legalcode". Accessed on July 13, 2016."

[4] "Public Domain License. Available at "https://creativecommons.org/publicdomain/zero/1.0/legalcode". Accessed on July 13, 2016."

[5] "GNU Free Documentation License. Available at "https://www.gnu.org/licenses/fdl-1.3.en.html". Accessed on July 13, 2016."

[6] "Free Art License. Available at "http://artlibre.org/licence/lal/en/". Accessed on July 13, 2016."

[7] "Understanding Metadata. Available at "http://www.niso.org/publications/press/UnderstandingMetadata.pdf", Technical Report, NSO Press, 2004. Accessed on July 13, 2016."

[8] "Stackoverflow FAQs: What is reputation? How do I earn (and lose) it? Available at "http://stackoverflow.com/help/whats-reputation". Accessed on July 14, 2016."

[9] "Twitter: An online social networking service. Available at "https://www.twitter.com". Accessed on July 14, 2016."

[10] "Facebook: An online social networking service. Available at "https://www.facebook.com". Accessed on July 14, 2016."

[11] B. T. Adler and L. D. Alfaro, "A content-driven reputation system for the Wikipedia," *Proceedings of the 16th international conference on World Wide Web*, 2007.

[12] "Open edX Platform for MOOC. Available at "https://open.edx.org". Accessed on July 13, 2016."

[13] "Github. Available at "https://github.com/". Accessed on 27 February 2015."

[14] "Apache Subversion. Available at "https://subversion.apache.org/". Accessed on 27 February 2015."

[15] "Concurrent Versions System. Available at "http://savannah.nongnu.org/projects/cvs". Accessed on 27 February 2015."

[16] "MediaWiki Open-Source Wiki Package. Available at "https://www.mediawiki.org/wiki/MediaWiki". Accessed on July 14, 2016."