International Journal of Distance Education Technologies
Volume 15 • Issue 2 • April-June 2017

Design of Open Content Social Learning based on the Activities of Learner and Similar Learners

Benneaser John, Department of Computer Applications, Karunya University, Coimbatore, India

J. Jayakumar, Karunya University, Coimbatore, India

V. Thavavel, Department of Computer Science, Prince Sultan Unviersity, Riyadh, Saudi Arabia

Muthukumar Arumugam, PSG Institute of Technology and Applied Research, Coimbatore, India

K. J. Poornaselvan, Department of Electrical and Electronics Engineering, Government College of Technology Coimbatore, Coimbatore, India

ABSTRACT

Teaching and learning are increasingly taking advantage of the rapid growth in Internet resources, open content, mobile technologies and social media platforms. However, due to the generally unstructured nature and overwhelming quantity of learning content, effective learning remains challenging. In an effort to close this gap, the authors designed and built an Open Content Social Learning (OCSL) system that compares different pedagogical strategies and algorithms intended to improve learning. Their results have shown increased effectiveness when recommending learning activities in a pedagogically appropriate order based on learning goals, historical learning preferences, and behaviors from other learners who had similar goals.

KEYWORDS

Collective Intelligence, Distance Education Technologies, Learner Centered Learning, Massive Open Online Courses, MOOC, Open Content, Open Pedagogy, Online Learning, Pedagogy, Social Network

1. INTRODUCTION

Computer-aided instruction (CAI) has evolved from its humble origins to the level of Massive Open Online Courses (MOOC). And the Internet and the entire World Wide Web (WWW) constitute the largest and most comprehensive knowledge base in the history of the world. Learners are living through an information explosion (Chiou & Shih, 2015). Rai & Chunrao (2015) states that "In recent years, MOOCs have attracted millions of learners around the world, through various MOOC providers, such as edX, Coursera, and Udacity. MOOCs allow millions of learners to enroll in courses form reputed universities around the world, such as Harvard University, Stanford University, Massachusetts Institute of Technology (MIT), and University of California at Berkeley (UCB). Outside of MOOCs, professors are creating and releasing their own content using tools such as Slideshare and YouTube." Every day, millions of learners make use of free, open online tools and resources (MacDonald, 2015) to create open learning content.

Open Educational Resources (OERs) are teaching and learning materials that anyone can use and share freely, without charge. Since first being coined by UNESCO in 2002, the term Open Educational Resources has evolved to meet the fast pace of the change and the diverse contexts in which it has now been used (Bossu, Bull, & Brown, 2012). The worldwide OER movement is rooted in the idea

DOI: 10.4018/IJDET.2017040105

of high quality education at no cost. The Cape Town Declaration (2007) states that "Educators worldwide are developing a vast pool of educational resources on the Internet, open and free for all to use. These educators are creating a world where each and every person on earth can access and contribute to the sum of all human knowledge. They are also planting the seeds of a new pedagogy where educators and learners create, shape and evolve knowledge together, deepening their skills and understanding as they go."

Open content learning resources such as MIT's OpenCourseWare project (OCW), TED videos, Khan Academy, YouTube videos, and the MERLOT(Malloy & Hanley 2001; Hanley 2015) project are a few examples of systems through which millions of learners learn on the web every day. However, the research on MOOCs shows that although thousands of people may register for a course, the number of students who complete the course successfully is generally much lower. Recent literature shows that although millions of people may register for MOOCs, completion rates vary from 0.7% to 52.1%, with a median value of 12.6% (Jordan, 2015). At Harvard University, the completion rate for the MOOC course CS50 is slightly below 1%. In contrast, 703 out of 706 students (99.6%) "Completed" CS50 on campus, the same course offered by the same lecturer (Parr, 2013). This is due to the lack of focus, engagement, motivation, and individual attention in those MOOC courses (Banerjee & Duflo, 2014; Rai & Chunrao, 2016). It's because of the lack of finding the right content, ability to collaborate with fellow learners, focus, and motivation. Hence, it is necessary both to provide the learners with appropriate information and knowledge resources and to make it easier for learners to engage, motivate, and collaborate with fellow learners.

Open learning enables learners to be self-determined and interest-guided. One of the challenges of open learning is, while the open content grows in popularity and we witness the proliferation of repositories and portals for Open Educational Resources(OER) content, it becomes more difficult for potential users to find the content and engage in learning (Caswell, Henson, Jensen, & Wiley, 2008). The power in OER is not in their production; it is in their reuse by other educators and learners. Due to the generally unstructured nature and overwhelming quantity of learning content, effective learning remains challenging. Learners are often unable to identify which material is needed, useful, and required at their level. Hence, open content learning design must assimilate the material from various sources and provide a new pedagogy that is appropriate to the needs of today's learners (Smyth, Bossu & Stagg, 2015). In this paper, we present our design for an Open Content Social Learning (OCSL) system that leverages Open Content to deliver an adaptive and personalized experience of the learners and similar learners and the need to recommend learning activities in a pedagogically effective order.

The remainder of this paper is organized as follows. Section 2 provides a literature review on different pedagogical strategies to make open online content learning effective. Section 3 covers the importance of Theories of learning and Open Pedagogy in making learning effective by leveraging Open Education Resources (OER). In Section 4, we discuss the architecture and implementation of our Open Content Social Learning (OCSL) System and the design of various modules that have been implemented. Section 5 covers the overall testing and the results of the research. We conclude the paper in Section 6 along with describing future work in this area.

2. RELATED RESEARCH

Several published research work were reviewed in the field of social learning and collective intelligence. Behaviorism, cognitivism, and constructivism are the three broad learning theories most often utilized in the creation of instructional environments (Siemens, 2014). These theories, however, were developed in a time when learning was not impacted through technology. Over the last twenty years, technology has reorganized how we live, how we communicate, and how we learn. Learning theories that describe learning principles and processes should be reflective of the underlying social environments.

Pedagogy is the art and science of teaching. Over the last decade numerous researches have been carried out about different pedagogical strategies to make the online learning environment effective. Ideas and information are being exchanged using multiple communication modes around the clock from anywhere in the world. A variety of research works are in progress for an effective instructional delivery strategy. Phillips et al. (2010) called it as a major challenge for instructional designers and practitioners for implementing authentic online learning to align the critical components of authentic tasks with effective learning principles. The rapid increase in online courses has definitely helped increase its reach but there is still a debate about the educational effectiveness of an online course. Kop et al. (2011) recommends an enhanced Pedagogy is required for online learning to be personalized based on learner's goals and style and compared with "learner like" learners (individualized and collaborative) as well as adaptive learning resources (organized and filtered) with the motivation and engagement tools. Learner's experiences with open learning do not always contribute to effective learning because some traditional pedagogical strategies are still being used. Over the past decade, researchers have investigated different pedagogical strategies for making the online learning environment effective. Sathiyamurthy & Geetha (2012) state that "The effectiveness of an e-learning system for distance education to a large extent depends on the relevancy and presentation of learning content to the learner". In a recent study, Kim & Reeves (2007) showed that the increase in online courses has definitely helped to reach millions of learners, but the educational effectiveness of online courses is a subject of debate. The goal of the adaptive presentation is to adapt the content to the user's goals, knowledge, and other relevant information. The architecture for an Adaptive Hypermedia System adapts the content of a hypermedia page to the user's goals, knowledge, preferences, and other user information for each individual user who is interacting with the system (Stern & Woolf, 2000). The effectiveness of integrating the pedagogies depends on high levels of interactivity among and between students and teachers and between students and the technologies that they use.

The World Wide Web with its current open and fragmented content is overwhelming large for any learner to embark upon the learning process with minimal search and absolute certainty. Valjataga et al. (2011) and Van Harmelen (2006) described, LMS have traditionally been mostly closed, leaving little room for learners to manage and maintain a learning space that facilitates their own learning activities as well as connections to peers and social networks across time and place. New techniques are being introduced, considering the World Wide Web as a LMS, for ensuring powerful learning experience. Nada Dabbagh and Anastasia Kitsantas (2012) discussed the idea of pedagogical approach for both integrating formal and informal learning using web content technologies.

Open content on the web can be found with some basic meta-data, such as the title, document type, and location, but additional metadata are required for the content to enable effective learning. Indexing, categorization and tagging methods are critical to filter the content to offer a personalized learning experience. Due to the massive amount and nature of the open web content, the learning experience process must be automated with efficient algorithms. The primary challenge of personalized results is presenting relevant content. An increasingly popular way to structure information is through the use of ontologies or graphs of concepts (Susan, Jason, Chaffee and Alexand, 2003). It is highly unlikely that the millions of users who have access to the Internet are so similar in their interests that one approach to browsing or searching, respectively, fits all needs. What is needed is a solution that will "personalize" the information selection and presentation for each user. Brusilovsky, Kobas & Nejdi (2007) suggest that students would be less likely to suffer from information overload if they were presented with personalized activities. Information overload is a concern due to the easy access to an abundance of online information sources (O'Donnell, Lawless, Sharp & Wade, 2015).

Another aspect of effective search and personalized results is consideration of the learner's profile. All learners are unique; no two will achieve the same learning outcomes across a range of subject areas. Clear guidance can be provided on the diverse learning needs of each student by collecting and continuously updating metadata that is stored for learners in user profiles. Chan (2000) describes that implicit profile creation based on observations of users actions has been used in more recent projects and describes the types of information that is available. This model considers the frequency of visits to a page, the amount of time spent on each page, how recently a page was visited, and whether the page was bookmarked. The user's learning behavior is used to create user profiles in several systems. Paireekreng & Wong (2010) observe that prior knowledge of each learner's activity and an effective user profile is required for personalization.

M.P. Cuéllar, M. Delgado, and M.C. Pegalajar (2011) have considered social networks to be a type of Learning Management System (LMS). Social Network Analysis (SNA) is conducted for teachers, learners, learning resources and their interactions. Vassileva, J. (2008) emphasizes that the two main goals of the design of social learning environments should be making them learner-centered and making learning more gratifying. In recent research, association rule-mining algorithms have been used to solve the problem of web page recommendations. A web usage log is used in adaptive association rule-based web mining, which attempts to personalize the results. Ujwala (2013) used the Apriori data-mining algorithm to generate association rules. The various recommendation approaches are context-aware approaches (Wang, Meng, & Zhang, 2012), semantic-based approaches (Di Noia, Mirizzi, Ostuni, Romito, & Zanker, 2012), cross domain-based approaches (Tang, Wu, Sun, & Su, 2012), peer-to-peer approaches (Kim, Kim,& Cho, 2008) and cross-lingual approaches (Schmidt, Scholl, Rensing, & Steinmetz, 2011). A new mining approach presented by Ujwala Wanaskar (2013) based on the combination of weighted association rule mining and text mining shows the best performance improvement compared to the existing methods.

Our specific research about increasing effectiveness and engagement leveraging Open content identified some gaps. Literature clearly supports two key ideas for effective learning.

a. Content personalization and relevancy.
b. Learner engagement and motivation.

Some of the gaps are analyzed during the literature review are listed below.

1. The open learning content volume is so high, and currently fragmented, it's very hard for learners to find the right content. This work made an initiative to address this gap with the idea of "Open Content Discovery". This will the increase the Content relevancy for learners.
2. Open content resources do not follow any meta-data standards. Currently, there's no method of easily searching across multiple Open Content repositories. This work is planning to address this gap with the idea of "Open Content Organize". Since, the volume of content is so high automated classification and taxonomy is important, and that's currently missing. This will the increase the Content personalization experience for learners.
3. As we studied in the literature, learners are not able to find the relevant content. There are deficiencies in the current methods of assessing the learner's engagement and efficiency. This research is focused on addressing the personalization of open content with the idea of "Open Content Personalization". There are two proposed ideas to increase the learner engagement and motivation.
    a. Content recommendation based on current learner's activity and goals.
    b.Personalized content recommendation based on similar learners and peer grouping.

3. OPEN PEDAGOGY AND THEORIES OF LEARNING

There are different variety of theories about how people learn. Fundamental theories of learning are still relevant to this digital age. Learning theories are conceptual frameworks describing how information is absorbed, processed, and retained during learning. Learning and Teaching methods are developed based on the learning theories.

Behaviorism is connected with behavior than with thinking, feeling, or knowing. This learning theory focuses on the objective and observable components of behavior. Behaviorism focuses on one particular view of learning: a change in external achieved through using reinforcement and repetition to shape behavior. Behavior theorists define learning as nothing more than the acquisition of new behavior based on external conditions.

Cognitive psychology is about identifying and describing mental processes that impact learning, thinking and behavior, and the conditions that influence those mental processes. This concept of mind as computer enabled several technology based ideas in learning & teaching:

1. **Intelligent Tutoring Systems:** Based on learner's responses to questions, these systems navigate the learner to the next lessons.
2. **Artificial Intelligence:** Software programs and algorithms to represent the mental processes used in human learning.
3. **Adaptive Learning:** Based on analysis of different kinds of cognitive activities, and required learning outcomes, this method of learning guides the learner through the learning process with the help of psychometric algorithms.

But, with this learning theory, as the learners learn new information is integrated with prior knowledge. This approach of learning was trying to fit human learning into the current restrictions of software programming.

Cognitive learning theory is deterministic, that behavior and learning are certain rules based and operate under predictable conditions. But, constructivists argue that learners learn based on their past experience and their present state. This means learning is continuous, and testing ideas through social contact and personal reflection. Thus knowledge is not just about content, but also values. The concurrence of constructivist approaches to learning and the development of the Internet has led to the development of a particular form of constructivist teaching, originally called computer-mediated communication (CMC), but which has developed into what Harasim (2012) now calls online collaborative learning theory (OCL). In the OCL theory, the teacher plays a key role not as a fellow-learner, but as the link to the knowledge community, or state of the art in that discipline. Learning is defined as conceptual change and is key to building knowledge. Learning activity needs to be informed and guided by the norms of the discipline and a discourse process that emphasizes conceptual learning and builds knowledge.

Connectivism is a relatively new theory, focuses on relationships between individual learning, the contributions of individuals to knowledge and its flow, and networks of learners. Downes (2014) sets out some design principles for Connectivism.

- Learner autonomy, in terms of choice of content and how they choose to learn
- Openness, in terms of access to the course, content, activities and methods of assessment
- Diversity: varied content, individual perspectives and multiple tools, especially for networking learners and creating opportunities for dialogue and discussion
- Interactivity: 'massive' communication between learners and co-operative learning, resulting in emergent knowledge

To achieve the desired learning outcomes, there must be some practices or style between the learning theories described above, and the design of learning systems. This strategy is commonly referred as a term Pedagogy. Open pedagogy is a set of teaching and learning practices possible in the context of the massive amount of open learning content. Traditional pedagogies imply the learning is static and content is closed and limited. Openness of content opens up limitless possibilities that inspire new perspectives and ideas. Learners are individuals and independent in the open learning process. Learners choose their own pace, their own direction, and their own connections. Learning design is focused less on specific outcomes or competencies than on process. It is about empowering learners to create real solutions to real problems.

Connectivism and open pedagogy are really the first attempt to radically re-examine the implications for learning of the Internet and the explosion of new communications technologies. In the following sections, we will discuss how the latest technologies and algorithms support, discovering, organizing and retrieving the massive numbers of learning content on the web.

Open pedagogy could be considered to be a blend of personalized adaptive design, algorithms and technologies, and networking among learners, which makes the learning process effective and engaging.

4. OPEN CONTENT SOCIAL LEARNING (OCSL) SYSTEM

In this section, we present our design for an Open Content Social Learning (OCSL) system that leverages Open Content to deliver an adaptive and personalized experience accounting for the pedagogical needs of the learners and similar learners and the need to recommend learning activities in a pedagogically effective order.

Structured content is content that has been broken down and classified using meta-data while unstructured content lacks most meta-data. Open Content on the web is mostly unstructured, that can be found with just some basic metadata like title, document type and location. Adding metadata to such open content is required to make the content suitable for effective learning. Indexing, categorization and tagging methods are critical to filter the content to offer personalized learning experience. Due to the massive size and nature of the open web content, efficient algorithms are required to add meta-data automatically.

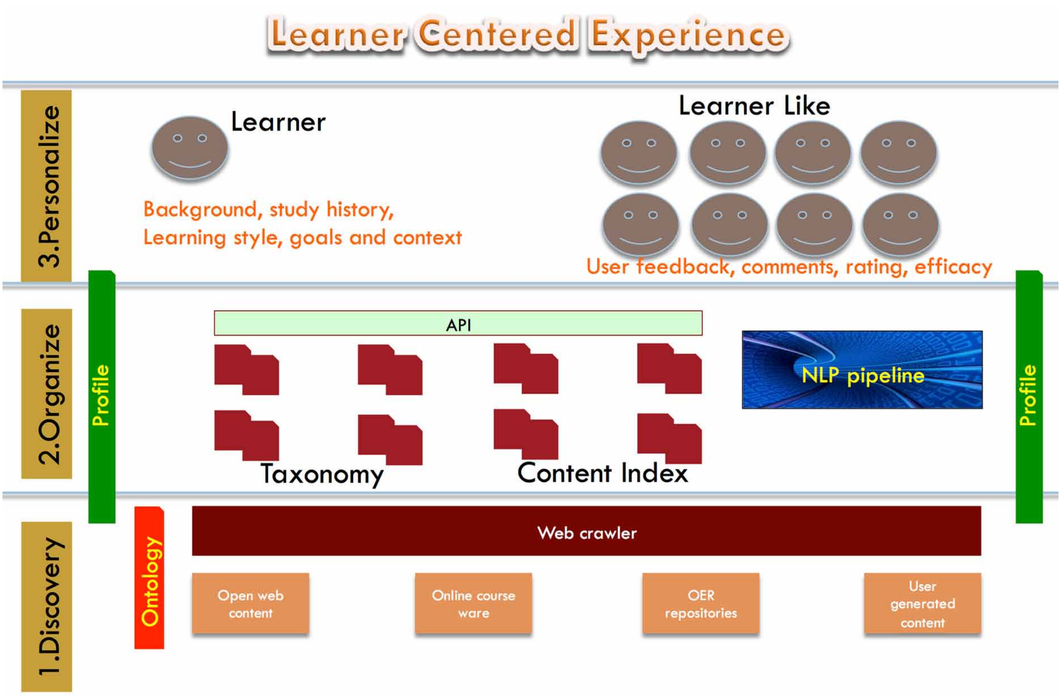
To achieve this, a three-layer architecture as shown in Figure 1 was identified for OCSL to improve pedagogical strategy and make the Learning effective:

Research shows that most of the Open Content learning platforms currently use standard search techniques by combining conventional information retrieval techniques that are based on page content, such as word vector space (Salton, & McGill, 1983), with link analysis techniques based on the hypertext structure of the Web, such as PageRank (Brin & Page, 1998) and HITS (Devi, Gupta, & Dixit, 2014).

Standard search techniques parse text into tokens to be indexed into an inverted index for any relevant information about documents (such as categories, subject or other attributes). Then the results are ranked to obtain an ordered list of results. The PageRank (Page, Brin, Motwani & Winograd, 1999) value for a page *u* is dependent on the PageRank values for each page *v* that is contained in the set B_u (the set that contains all of the pages that link to page *u*), divided by the number L(v) of links from page *v*. The PageRank value for any page *u* can be expressed as

$$PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)}$$

Figure 1. Overview of the Learner-Centered Learning Experience leveraging Open Content



The PageRank algorithm (Brin & Page, 1998) attempts to provide an objective estimate of the Web page importance. However, the importance of the Web pages is subjective for different users. The true relevancy of a page depends on the interests, goals and existing knowledge of the individual users; a global ranking of a Web page might not necessarily capture the importance of a page for a given individual user.

OCSL expands the scope of the search to generate more personalized results and greater learning engagement. Since the system needs to handle large volume of learning data and automate the categorization and indexing, it requires both offline process and online process. The modules in open discovery and organize module are executed in offline but are updated periodically, whereas the modules in open content personalization is executed online. Following are the high level OCSL modules.

1. Offline Process:
    a. Content discovery and indexing using the open source projects Apache Nutch and Apache Solr
    b. Content clustering and classification to populate content attributes (meta-data)
2. Online Process:
    a. Dynamic pedagogical engine that personalizes the learning experience and content ranking based on the learner profile attributes and content attributes
    b. Dynamic pedagogical engine that personalizes the learning experience and content ranking based on the learner profile attributes, content attributes and similar learner profile attributes

Each module performs its defined function and exchanges information with other modules, as shown in Figure 2 and Figure 3. The functions of these modules are discussed individually below.

4.1. Macro Algorithm: Offline Process

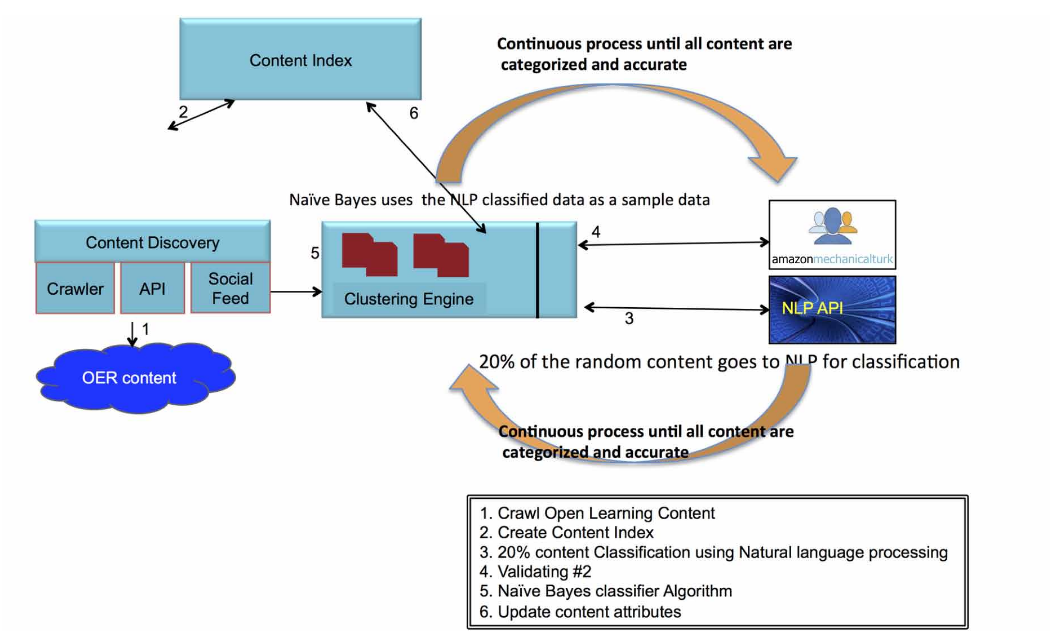
Step 1: Content is consumed using web crawler, API calls and streaming social data feed from various OER repositories and then send to cluster engine for categorization. Apache Solr tool is used as a web crawler.
Step 2: The clustering engine consumes the discovered open content and sends 20% of content to Natural Language Processing(NLP) API. The remaining 80% is sent to classifier which classifies using the Naïve Bayes classifier algorithm.
Step 3: NLP API classifies 20% of the content.
Step 4: Categorized open content are sent to Amazon Mechanical trunk for testing to make sure the content classified is accurate. If not, the feedback goes to the classification to update.
Step 5: Naive Bayes algorithm classifies 80% of the content using the 20% of sample content classified by NLP API.
Step 6: Once the content is correctly classified with attributes (meta-data), it is loaded into the content index.
Step 7: The content index indexes the attributes and stores it inside the Apache Solr container. This content index is updated periodically.
Figure 2 below summarizes the above offline process steps.

4.2. Macro Algorithm: Online Process

Step 1: The learner enters goals, topics, interest etc. as an input to the query formulator to deliver personalized open content.
Step 2: The query formulator uses three different algorithms to search the open content from context index as follows:

- **Content Hierarchy and Learner Attribute based Matching (CHLAM):** This algorithm for the first time in the literature attempts to search open content utilizing the learner attribute (profile data) stored in profile repository and content attributes.
- **Content Hierarchy and Similar Learner Attribute based Matching (CHSLAM):** In revision to the above, this algorithm also for the first time in the literature attempts to search

Figure 2. Offline process: Architecture and design of the OCSL system



open content utilizing the learner attributes (profile data) stored in profile repository and content attributes along with similar learner activities.

- **Conventional Content Search:** This algorithm is devised for comparison purpose. This algorithm uses the Vector Space Model of Information Retrieval as described by Salton and McGill (1983). Here's the logic of this algorithm: More times a query term appears in a document relative to the number of times the term appears in all the documents in the collection, the more relevant that document is to the query.

Figure 3 below summarizes the above online process steps.

4.3. Content Discovery and Index

The role of content discovery is to crawl open content from the Internet, i.e., the World Wide Web and social media, and to locate content to present to the user. The content discovery referred in Figure 2 is configured to collect content from three sources: 1. Crawling OER content sites 2. Streaming API against social media platforms 3. API calls against learning platforms such as MERLOT (Hanley, 2015), OER Commons, Gooru learning.

The following are some of the Open Content sources that the OCSL system ingests content from for the purposes of this research:

- OER Commons API
- Merlot API
- Gooru learning API
- https://www.facebook.com/openeducationeuropa_
- https://www.facebook.com/OEConsortium_
- https://www.facebook.com/pages/OER-Africa/_
- https://twitter.com/OERCommons_
- https://api.twitter.com/1.1/search/tweets.json?q=oer_
- http://www.coursetalk.com_

Figure 3. Online process: Architecture and design of the OCSL system

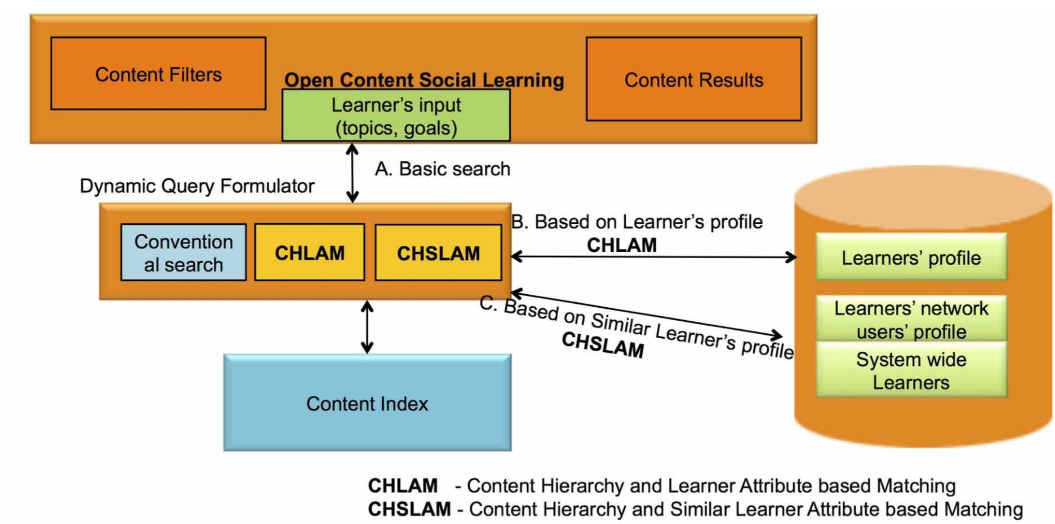


Figure 4. Content Repository Schema for capturing meta-data information

File Path ▾ : ~/Documents/phd/code/osl-v2-code/solr core/solr/conf/schema.xml

schema.xml (no symbol selected)

```xml
    <fields>
        <field name="id" type="string" stored="true" indexed="true"/>
        <!-- core fields -->
        <field name="batchId" type="string" stored="true" indexed="false"/>
        <field name="digest" type="string" stored="true" indexed="false"/>
        <field name="boost" type="float" stored="true" indexed="false"/>
        <!-- fields for index-basic plugin -->
        <field name="host" type="url" stored="false" indexed="true"/>
        <field name="url" type="url" stored="true" indexed="true" required="true"/>
        <field name="content" type="text" stored="true" indexed="true"/>
        <field name="title" type="text" stored="true" indexed="true"/>
        <field name="cache" type="string" stored="true" indexed="false"/>
        <field name="tstamp" type="date" stored="true" indexed="false"/>
        <field name="anchor" type="string" stored="true" indexed="true" multiValued="true"/>
        <field name="type" type="string" stored="true" indexed="true" multiValued="true"/>

        <field name="mainType" type="string" stored="true" indexed="true" multiValued="false"/>
        <field name="contentLength" type="long" stored="true" indexed="false"/>
        <field name="lastModified" type="date" stored="true" indexed="false"/>
        <field name="date" type="date" stored="true" indexed="true"/>

        <!-- fields for languageidentifier plugin -->
        <field name="lang" type="string" stored="true" indexed="true"/>
        <!-- fields for subcollection plugin -->
        <field name="subcollection" type="string" stored="true" indexed="true" multiValued="true"/>
        <!-- fields for feed plugin (tag is also used by microformats-reltag)-->
        <field name="author" type="string" stored="true" indexed="true"/>
        <field name="tag" type="string" stored="true" indexed="true" multiValued="true"/>
        <field name="feed" type="string" stored="true" indexed="true"/>
        <field name="publishedDate" type="date" stored="true" indexed="true"/>
        <field name="updatedDate" type="date" stored="true" indexed="true"/>
        <!-- fields for creativecommons plugin -->
        <field name="cc" type="string" stored="true" indexed="true" multiValued="true"/>
        <!-- fields for tld plugin -->
        <field name="tld" type="string" stored="false" indexed="false"/>
        <field name="technicalFormat" type="string" stored="true" indexed="true"/>
        <field name="userRating" type="float" stored="true" indexed="true"/>
        <field name="category" type="string" stored="true" indexed="true" multiValued="true"/>

    </fields>
```

Line 65 Col 9 | XML | Unicode (UTF-8) | Unix (LF) | Last saved: 10/18/15, 8:24:58 PM | 5,291 / 554 / 91

Open Source Projects Apache Nutch 2.0 and Apache Solr 3.6 have been used for the crawler and content indexing. Solr indexes the attributes and stores the content as a repository. This repository is created and updated periodically in an offline process that involves the content discovery process mentioned above. Figure 4 below shows the schema that is used for the content store.

4.4. Clustering and Classification Engine

Content clustering entails grouping similar uncategorized documents together based on similarity measures. Content classification categorizes and organizes content by combining multiple methods of context-sensitive analysis. Thus, the purpose of this engine is to process the content from various sources to properly map it to the taxonomy that we generated to support STEM (science, technology, engineering, and mathematics) content. Although this system is designed as a generic solution, the first completely implemented and tested version was with STEM content.

The clustering engine referred in Figure 2 consumes content from multiple sources (Nutch Crawler, Federated API search, and Streaming API for social media feeds) and performs the following steps:

1. Alchemy's machine learning APIs (Quercia, Askham, & Crowcroft, 2012) are used for categorizing the content. OCSL uses the Taxonomy API to perform classification. The Entity API calls fetch the desired Internet web page, normalizes it, and extracts named entities, topics, and other content.

a. http://www.alchemyapi.com/api/taxonomy_calls/urls.html
b. http://www.alchemyapi.com/api/entity/urls.html#rurl

Using the Taxonomy and Entity API, content metadata is updated in the Solr content repository.

2. As recommended by Wang, Kraska, Franklin, & Feng (2012), OCSL leveraged a hybrid human-machine approach in which machines are used to perform an initial, coarse pass over all of the data, and people are used to verify only the most likely matching pairs. OCSL integrates with the Amazon Mechanical Turk API to verify the classified content.
3. Using the Apache Mahout framework and Naive Bayes classifier algorithm (Patil & Pawar 2012), OCSL automatically classifies documents using a training set developed from the previous two steps. The training set includes documents that are already associated with a category. Using this set, the classifier determines, for each word, the probability that it reflects a document that belongs to each of the considered categories. To compute the probability that a document belongs to a category, the classifier multiplies together the individual probabilities of having each of its words in this category. The category that has the highest probability is the category that the document is most likely to belong to.
4. OSCL updates the content index engine with all of the taxonomy attributes (URL, content category, content sub category, content type, last modified, and many more).

4.5. Dynamic Query Formulator (Pedagogical Engine)

The Dynamic Query Formulator referred in Figure 3 is the core component of the OCSL system design. The pedagogical engine uses a dynamic query formulator algorithm that was developed through this research to navigate a learner's learning experience by analyzing his/her user interactions and prior learning knowledge on any given topic. The OCSL pedagogical engine also dynamically generates a query based on similar learners' learning experiences.

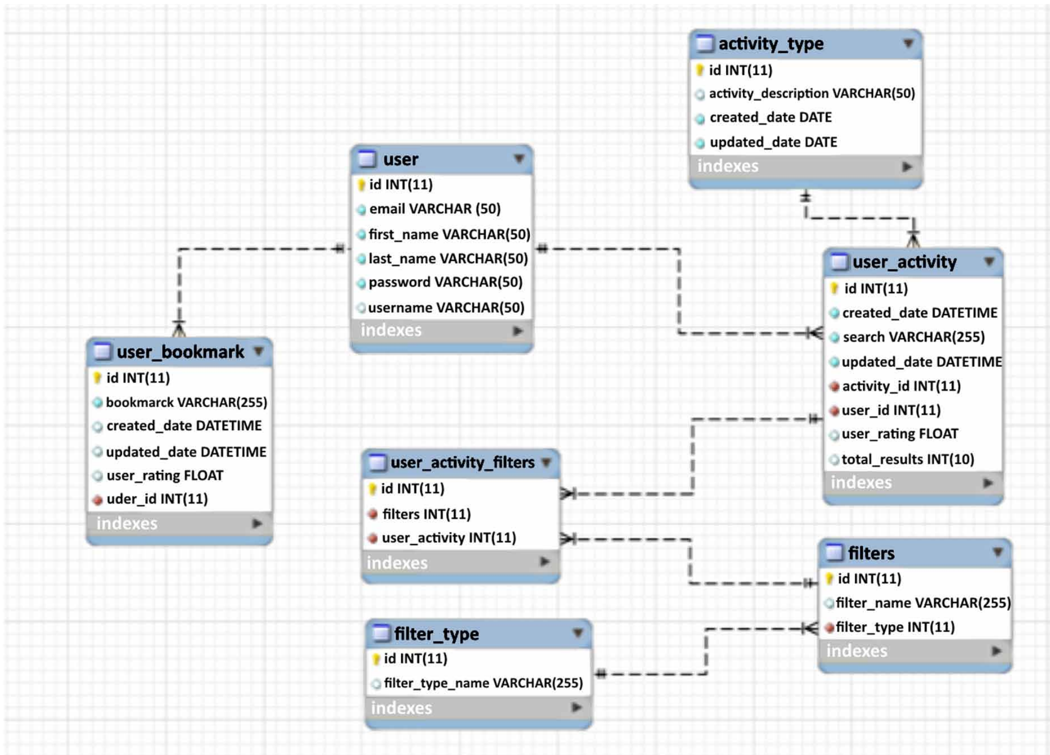
4.6. Content Hiearchy and Learner Attribute-based Matching (CHLAM)

Content Hierarchy Learner Attribute-based Matching (CHLAM) enhances the conventional search experience by building a user profile to provide more personalized search results based on learning style, type of content, recent activity, content categories, or other interests of the users. Although Solr's default similarity calculation works well on generic text, OCSL significantly improved the relevancy of the search results by passing along additional information to Solr. Additionally, OCSL passed user behavior such as past clicks or favorite lessons, OCSL passes this information along to Solr so that it can better understand which documents are related to each other based upon the similar users acted upon together. This is the core logic of CHLAM algorithm.

To build an intelligent pedagogical learning engine based on attributes, this system ensures that both users and documents are tagged with the same types of attributes. As an enhancement, we are implicitly and explicitly collecting information from learners about their learning behaviors, learning goals, and other criteria. Basically, the pedagogical engine is responsible for figuring out both the most appropriate way to construct the queries and which data to use in them to optimize the relevancy of the learner's learning experience. While a conventional search engine builds a sparse matrix of terms that are mapped to documents in the content index, OCSL enhances the design to map the user's behavior to those documents. When a conventional search engine receives documents with fields, it parses the content from the fields into tokens in an inverted index. Our behavior-based pedagogical engine matches metadata on user behavior with those tokens to filter the search results.

Figure 5 below shows the data model that was developed to build this system. The Personal information was implemented as a set of attributes, which store static personal characteristics about the learner, for example, username, password, unique ID, activities, and e-mail. The system associates the learner's knowledge level with each concept of the domain model. It then continuously updates

Figure 5. Open content social learning system data model



the skill level of the learner, developing a map of the learner's state of knowledge, to support the personalization of his/her learning experience.

The Learner Attribute-based Search enables the system to classify users and content into a hierarchy that goes from more general to more specific categories, but it is further possible to query this hierarchy and apply a stronger relevancy weight to more specific matches. The Learner Attribute-based Search enables the system to generate search based on the personalized attributes for the users. The end result is that by using query weights on terms that combine a measure of their probability (most likely to least likely) and their specificity (most descriptive to least descriptive), a fuzzy query can be constructed to match documents that match any of the criteria; at the same time, it boosts documents to the top of the search results that match the best combinations of those attributes within the hierarchy.

The query parameter also allows the system to weight the fields differently. This parameter can be used to make a query match in one field more significant than a query match in another field.

$$qf = field_1^{v1} + field_2^{v2} + \ldots + field_n^{vn}$$

where qf is the Query Fields, and v is the weight for each attribute, based on the learner's goals and interests as calculated and applied dynamically. In our approach, we personalize PageRank scores by assigning weights to the fields based on matched goals and activities based on the learner. Each learner will have a personalized weight for each attribute.

4.7. Content Hiearchy and Similar Learner Attribute-based Matching (CHSLAM)

By mapping the learning behavior of users to documents, we are effectively creating links in the index between documents. Klašnja-Milicevic, et al. (2011) recommended that similar users learn similar content, which means that documents that are mapped to similar users are likely related. To make use of these relationships to recommend learning items to a new user, we find other similar users and recommend other items. OCSL provides a mechanism to form a social network among the learners who have similar learning interests, preferences, and learning experiences based on the data collected. A learning group in OCSL is a group of learners who share common learning goals and mutually recommend learning content that meet those goals. OCSL uses User-based Collaborative filtering and Item-based Collaborative filtering to filter the learning content and recommend learning activities in a pedagogically effective order. Our algorithm, called CHLAM adds weight dynamically to the search for an effective personalized search based on the learners' past activity. CHSLAM algorithm is an attempt to improve the performance of CHLAM algorithm by adjusting the weight based on the learner and similar learners. The similarity users are measured by all the users who have similar profile properties. Similar users based on activities or rating information of like-minded users (called neighbors) has been widely adopted. However, there is still considerable room for improving the quality of recommendation. Similarity functions in traditional Collaborative Filtering(CF) gives an equal weight to each of the activity involved by both users. For example, when a learner clicks an assignment vs completes a quiz both gives a same weight. The proposed algorithm gives a different weight based on the type of activity. In the new similarity function, the rating of a user on an item is weighted by the item similarity between the item and the target item. Following is how CHSLAM algorithm works by adjusting the weight based on the learner and similar learners.

1. Find similar learners and group them as a peer group. Following is the algorithm to find similar learners.

Step 1: Group the users with similar activities.

FROM users_activity JOIN users_activity.id = users_activity.id WHERE users_activity.user_id = current_user AND users_activity.user_id != current_user GROUP BY users_activity.user_id

Step 2: Construct data as a list of the User ID, Activity ID and count of the activities. A sample data would look like the following.

User ID, Activity ID, Count
9101,1001,18
9101,1002,11
9101,1003,6

Step 3: Input this data set to calculate peer group of the learners based on the activity count and compare with other profile attributes like goals, last activity date, lessons etc.

2. Identify all the most commonly used attributes across the users in the peer group.
3. Calculate the median value for each attribute.

5. EVALUATION

To evaluate our design, we conducted a Web crawl against Open Content Resources (OCR) and implemented a dynamic query formulator engine. We performed an experimental study that focused on STEM engineering students. Our study explored the results of the following three algorithms, to validate the idea of effective learning by personalizing the content results. The study lasted for almost three months. Learners were grouped into 15 groups. These learners were focusing on learning STEM content. Approximately, 18,000 learning content was classified and presented to the learners.

Around 300 learners were grouped into 15 groups, and they used the OCSL system over the period of three months.

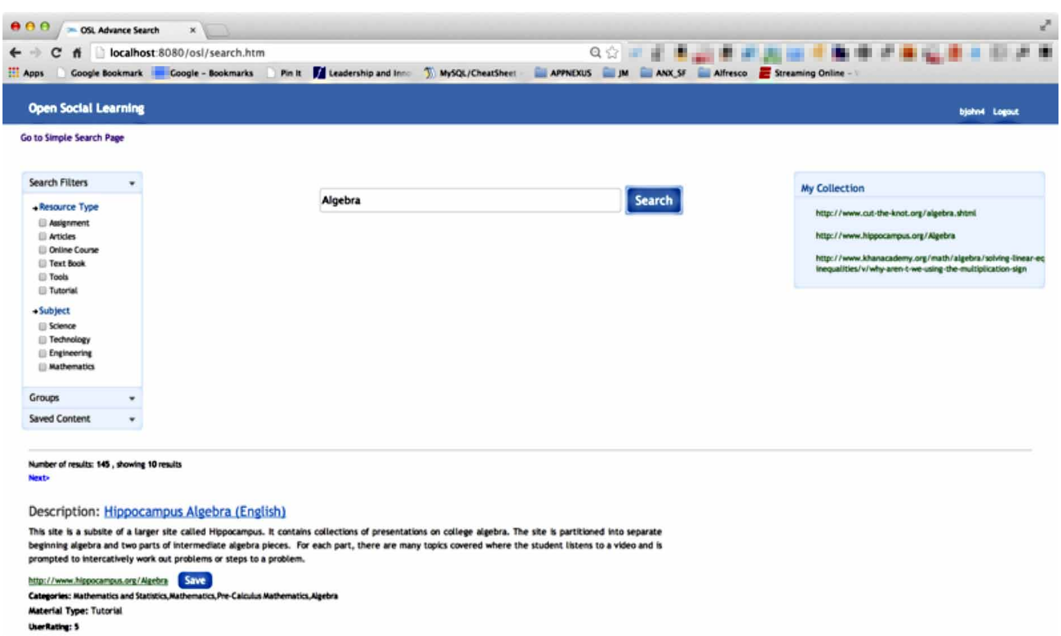
1. **Algorithm 1:** Basic search using inverted index and page ranking conventional algorithm
2. **Algorithm 2:** Search based on the Content Hierarchy and Learner Attribute-based Matching (CHLAM) of the OCSL system
3. **Algorithm 3:** Search based on CHLAM and Similar Learners Attribute-based Matching (CHSLAM) of the OCSL system

We asked each learner to use our OCSL system after they entered their goals and profiles into our system. We did not provide any information about the main goal of the system. The learners were expected to use the platform and learn based on their choice of preferences. A results page was shown with the recommended content based on the three different types of algorithms mentioned above. Figure 6 is a screen shot of the OCSL system.

5.1. Testing Approach

Comparing search results and recommendation systems is difficult. The best way to experiment with different relevancy parameters is to run A/B experiments that randomly divide users into groups over the same time period, with each group interacting with a different algorithm. We placed 1/3 of the users into a control group that represents the conventional search algorithm, 1/3 of the users into a test group for algorithm CHLAM of OCSL, and 1/3 of the users into a test group for algorithm CHSLAM of OCSL. We hoped this approach would allowed us to measure each of the three groups independently to determine which group performed best by statistically validating the results. However, we found that this approach was limited in that it did not allow us to test multiple algorithms with the same learners.

Figure 6. OCSL System screen shot



Beyond A/B testing, another common method for measuring the relative performance of algorithms involves generating test data and performing comparative analysis using the generated log data (Khosla, & Bhojane, 2013). To experiment with learning activities in detail, behavioral patterns were extracted from the log files and user activity database table.

There are two aspects of a search result set that determine the quality of the results, the precision and recall, as Powers and David (Powers & David, 2011) suggest. Precision is the fraction of the retrieved documents that are relevant. A precision of 1.0 means that every result that is returned by the search is relevant, but there could be other relevant documents that were not a part of the search result.

$$precision = \frac{\left|\left\{relevant\ documents\right\} \cap \left\{retrieved\ documents\right\}\right|}{\left|\left\{retrieved\ documents\right\}\right|}$$

Recall is the fraction of the relevant documents that are retrieved. A recall of 1.0 means that all of the relevant documents were retrieved by the search, irrespective of the irrelevant documents also included in the result set.

$$recall = \frac{\left|\left\{relevant\ documents\right\} \cap \left\{retrieved\ documents\right\}\right|}{\left|\left\{relevant\ documents\right\}\right|}$$

If all of the documents are retrieved, then the recall is perfect but the precision may not be good. On the other hand, if the document set contains only a single relevant document and that relevant document is retrieved in the search, then the precision is perfect but again the result set will not be good. This relationship shows a trade-off between the precision and recall, in which they are inversely related. Because the web has an enormous collection of documents, it makes sense to provide a few relevant and good hits as opposed to adding irrelevant results in the result set. The F-score is a measure of a test's accuracy. It considers both the precision p and the recall r of the test to compute the score:

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

where F1 is the F-Score.

In this approach, we can take previously saved user behavior data from log files and test how well each of the candidate algorithms predicts the results that were previously acted on by the users. In the case of OCSL, we take the list of search results for every search or recommendation run for the user and plot them in aggregate on a precision versus recall graph, showing whether the algorithm made the correct prediction based on the user's historical behavior. For example, the correct prediction might be defined in terms of which learning materials a user consumed, and thus, any query model that resulted in higher precision and recall for that learning content would be considered to be a better algorithm.

5.2. Results

This section describes the testing model and different experiments to test multiple algorithms. Comparing search results and content effectiveness is difficult. One way to experiment with different relevancy parameters is to run A/B experiments that randomly divide users into groups over the same time period, with each group interacting with a different algorithm. However, this research found that this approach was limited in that it did not allow us to test multiple algorithms with the same learners. Beyond A/B testing, another common method for measuring the relative performance of algorithms involves generating test data and performing comparative analysis using the generated log. To experiment with learning activities in detail, behavioral patterns were extracted from the log files and user activity database table.

Efforts were taken to analyze the log files and determine the following metrics, precision, Recall and F-score based on learner activity for each algorithm discussed in this work and are depicted in Figure 7, 8 and 9. The F-Score can be interpreted as a weighted average of the precision and recall, where an F-Score reaches its best value at 1 and worst at 0. The F-score shows an absolute score for an algorithm that strives for good balance between the precision and recall.

We measured both median and mean values. Table 1 above listed all the values. Since the sample size is large and doesn't include outliers, the mean value usually provides a better accuracy. The F-Score can be interpreted as a weighted average of the precision and recall, where an F-Score reaches its best value at 1 and worst at 0. The average F-Score value for conventional algorithm was 0.0035, and for CHLAM algorithm it was 0.0191 and for CHSLAM algorithm it was 0.0202. Based on the tests, CHSLAM algorithm yielded better F-Score results. To obtain a subjective evaluation of the OCSL system, we organized a non-mandatory questionnaire that collected information on learners with respect to the main features of the system. More than 65% of the learners reported that the system recommended personalized results and was able to focus on the correct content. Based on the retrieved data by OCSL, learner engagement metrics were calculated during this research. Based on the learners' activity and the involvement during the learning period, learners' leveraged the CHSLAM algorithms engaged better. All 100% of the learners using conventional algorithm engaged only 25% or below. 88% of learners leveraged CHLAM algorithm engaged only 25% or below and remaining 12% engaged 50% or below. This is better than conventional algorithm. But 4%

Figure 7. Precision Metric of Conventional, CHLAM &CHSLAM algorithms in OCSL

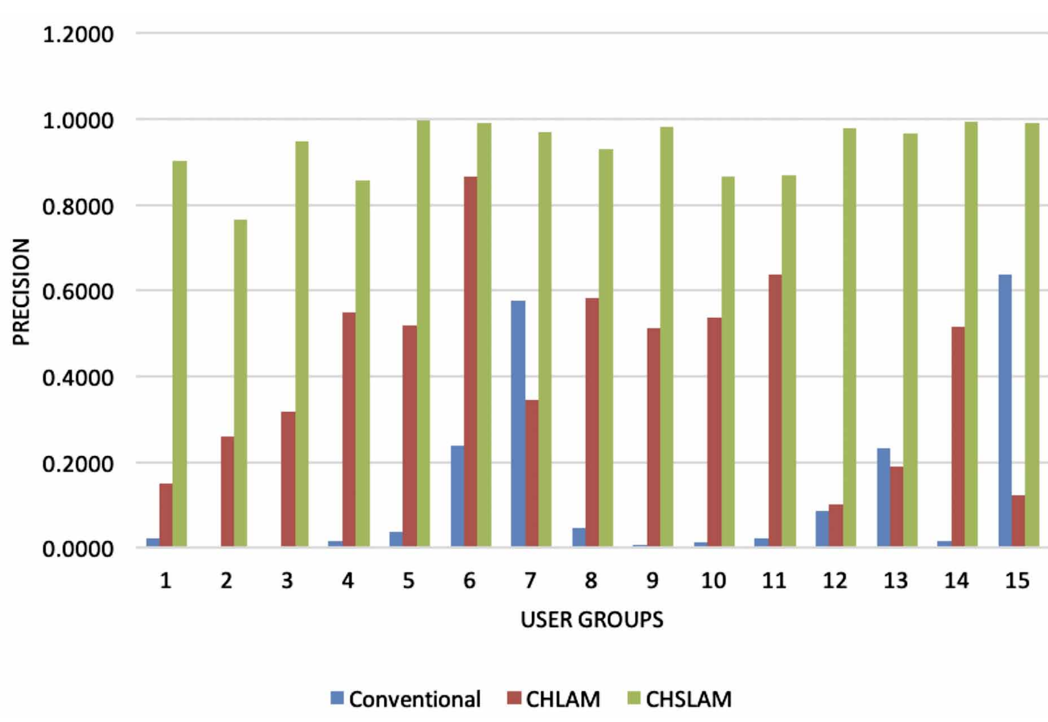


Figure 8. Recall Metric of Conventional, CHLAM and CHSLAM algorithms in OCSL

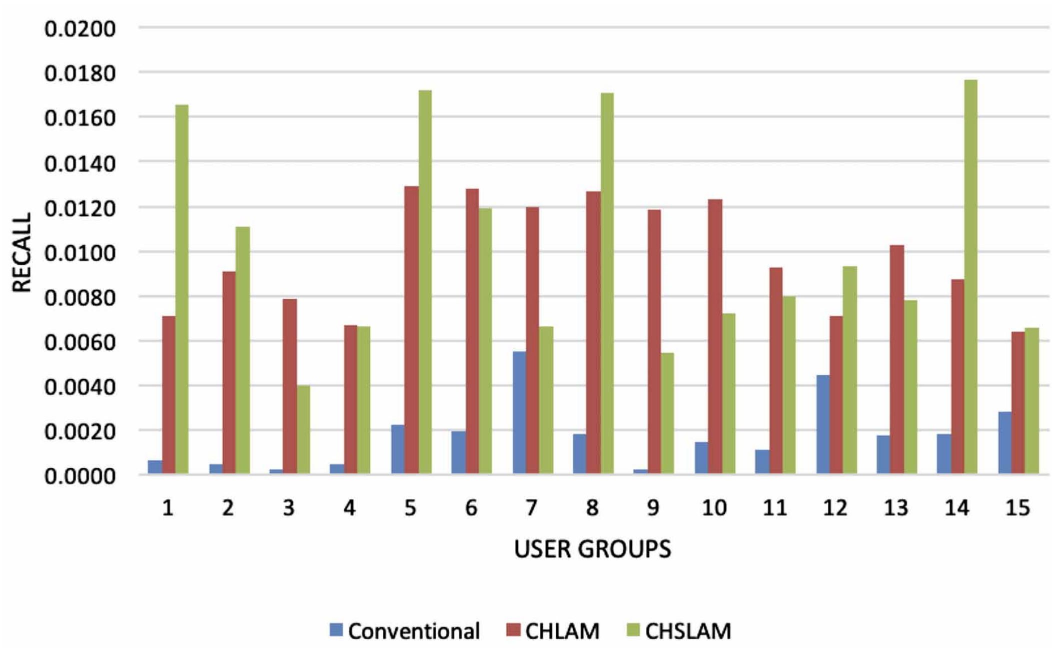
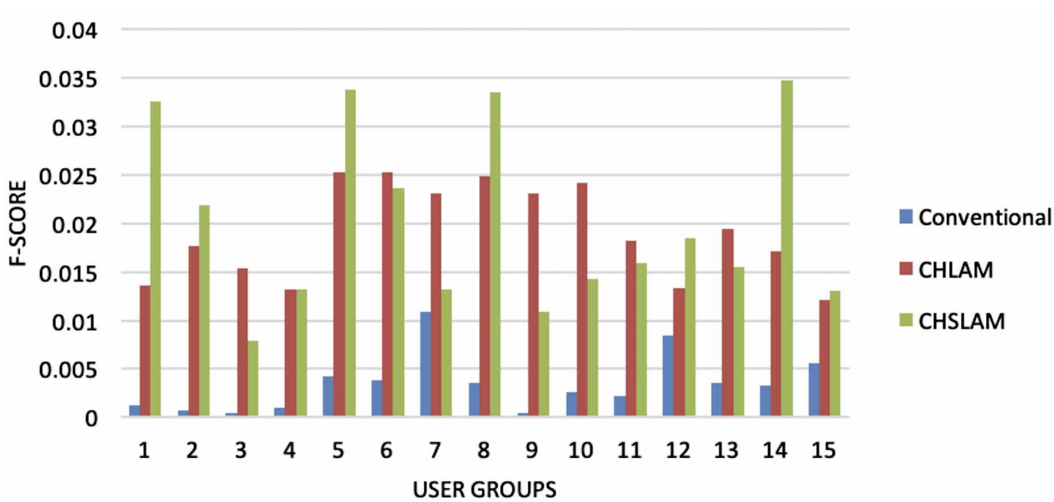


Figure 9. F-Score Metric of Conventional, CHLAM and CHSLAM algorithms in OCSL



of learners leveraged CHSLAM algorithm engaged 75% or below, and 34% learners engaged 50% below and 62% of learners engaged 25% or below.

This shows CHSLAM algorithm helps learner engage better than the learners who leverage conventional or CHLAM. Figure 10 below shows the graphincal representation of the above details.

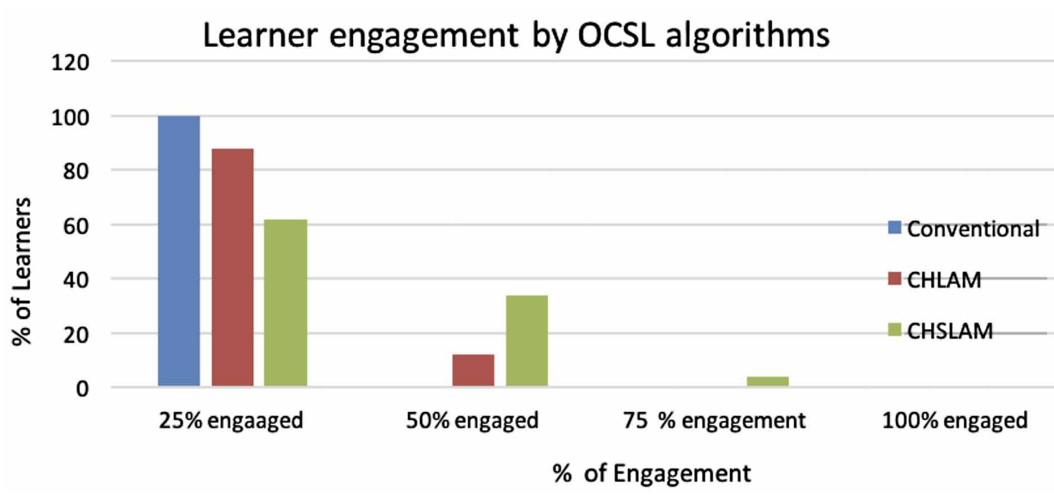
Table 1. Mean and median values of Precision, Recall and F-Score metrics based on Conventional, CHLAM and CHSLAM algorithms

Precision	Conventional	CHLAM	CHSLAM
Mean	0.1315	0.4140	0.9334
Median	0.0241	0.5134	0.9658

Recall	Conventional	CHLAM	CHSLAM
Mean	0.0018	0.0098	0.0102
Median	0.0018	0.0093	0.0080

F-Score	Conventional	CHLAM	CHSLAM
Mean	0.0035	0.0191	0.0202
Median	0.0033	0.0182	0.0158

Figure 10. Learner engagement comparison between conventional, CHLAM and CHSLAM algorithms



6. CONCLUSION AND FUTURE STUDY

This research work started with the motivation of increasing effectiveness of open learning, We began with a review of the existing OER search engines and studied the literature that pertain to the effectiveness of Open content searches. We found that most of the existing research measured effectiveness is based on surveys and real-time user metrics was not considered. We designed and implemented an end-to-end home grown system by leveraging open source technologies and existing content repositories. Research hypothesis was, fragmented open content ecosystem making it harder for learners to find the right content. Learners spend more time finding the right content than making use of the content.

OCSL system implemented some of the unique features during this research to validate the hypothesis.

1. To effectively categorize and search the content, OCSL required to consume all content and meta-data part of the system. So, we implemented a system to handle federated search across heterogeneous systems supporting both real time and batch data load.
2. Due to the size and nature of the open web content, automation of managing taxonomy is a key element to make the end-to-end process very effective. OCSL applied both supervised learning and unsupervised learning algorithms to categorize the content.
3. During the content indexing process, index time boosting/ranking was applied to the content.
4. Learners can interact with the system by inputting their goals, topics of interest etc. Implicit profile creation based on observations of learners' activities has been used in the OCSL system. The enhanced Pedagogical Engine is designed to support recommending the learning content based on the learner's goals, past activity. Query time boosts are applied based on the content attributes that matters the most to the current learner. We come up with a new algorithm named CHLAM to personalize the learning experience.
5. To further increase the engagement of learner, as research suggested we made an attempt to personalize learning experience based on similar learners' activities on the OCSL system. The system created peer groups based on the similar activities, and derived the content attribute weight to apply during the search time boosting.
6. We engaged several groups of learners to use the OCSL system and collected data both from the log files as well as database. We used Precision, Recall and F-Score as well as the activities of the learner to measure the effectiveness and engagement.

It is validated by spending less time on searching content, and consuming relevant content, learning process was effective. Learners engaged more successfully based on the CHSLAM algorithm compared to the CHLAM and conventional algorithms.

There's further research possible by extending the personalized mechanism and pedagogical aspects of OCSL to increase the engagement of learners by having the influencers and mentors interact with the peer group. Graph database technologies can be used to map the learners and mentors to manage the relationships effectively. OCSL system can be further improved by enabling learners with all the social networking functionality. Graph database technologies support building social learning graph, and then we can build models of influence based on the learners' activities. Current system doesn't expose learner's activities to other learners. It uses only in the background to identify similar learners in the peer grouping concept. But the idea is that when learner sees their peers performing an action such as taking a lesson or answering to questions, that learner influenced to perform similar action. Also, when learners perform similar action, it strengthens the peer group. Social Learning Theory explains about people learning new ideas and develop new behaviors by observing other people. It is suggested to make further progress on this research by exposing the similar learner activities to the learners, in the social learning environment.

REFERENCES

Antunes, C., & Oliveira, A. (2002, July). Using context-free grammars to constrain apriori-based algorithms for mining temporal association rules.*Proc. of the Workshop on Temporal Data Mining*.

Banerjee, A. V., & Duflo, E. (2014). (Dis) organization and Success in an Economics MOOC. *The American Economic Review*, *104*(5), 514–518. doi:10.1257/aer.104.5.514 PMID:25214652

Bansal, T., Chabra, S., & Joshi, D. (2013). Current Initiatives and Challenges to OERs in Indian Higher Education.

Barrett, R., Maglio, P. P., & Kellem, D. C. (1997, March). How to personalize the Web.*Proceedings of the ACM SIGCHI Conference on Human factors in computing syls* (pp. 75-82). ACM. doi:10.1145/258549.258595

Bledsoe, T. S., Harmeyer, D., & Wu, S. F. (2014). Utilizing Twitter and# Hashtags Toward Enhancing Student Learning in an Online Course Environment. *International Journal of Distance Education Technologies*, *12*(3), 75–83. doi:10.4018/ijdet.2014070106

Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, *56*(18), 3825–3833. doi:10.1016/j.comnet.2012.10.007

Brusilovsky, P., Kobsa, A., & Nejdl, W. (2007). *The adaptive web: methods and strategies of web personalization* (Vol. 4321). Springer Science & Business Media. doi:10.1007/978-3-540-72079-9

Chan, P. K. (2000). Constructing web user profiles: a non-invasive learning approach. In *Web usage analysis and user profiling* (pp. 39–55). Springer Berlin Heidelberg. doi:10.1007/3-540-44934-5_3

Cheung, B. S., Hui, L. C. K., Yiu, S. M., Lee, J. K., Kwok, S. L., & Leung, K. C. (2003). Content Engineering Agent: A TBL-Based E-Course Development Tool with TQM. *International Journal of Distance Education Technologies*, *1*(2), 57–71. doi:10.4018/jdet.2003040104

Cheung, K. S., Lam, J., Im, T., Szeto, R., & Yau, J. (2008, December). Exploring a pedagogy-driven approach to e-courses development. In *International Workshop on Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing* (Vol. 1, pp. 22-25). IEEE. doi:10.1109/ETTandGRS.2008.267

Chiou, Y., & Shih, T. K. (2015). Auto Grouping and Peer Grading System in Massive Open Online Course (MOOC). *International Journal of Distance Education Technologies*, *13*(3), 25–43. doi:10.4018/IJDET.2015070102

Cuéllar, M. P., Delgado, M., & Pegalajar, M. C. (2011). Improving learning management through semantic web and social networks in e-learning environments. *Expert Systems with Applications*, *38*(4), 4181–4189. doi:10.1016/j.eswa.2010.09.080

D'Antoni, S. (2009). Open Educational Resources: reviewing initiatives and issues. *Open Learning: The Journal of Open, Distance and e-Learning*, 24(1), 3-10.

Dabbagh, N., & Kitsantas, A. (2012, January). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education*, *15*(1), 3–8. doi:10.1016/j.iheduc.2011.06.002

de Waard, I., Koutropoulos, A., Keskin, N., Abajian, S. C., Hogue, R., Rodriguez, O., & Gallagher, M. S. (2011). Exploring the MOOC format as a pedagogical approach for mLearning. Proceedings from mLearn.

Devi, P., Gupta, A., & Dixit, A. (2014). Comparative Study of HITS and PageRank Link based Ranking Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, *3*(2).

Di Noia, T., Mirizzi, R., Ostuni, V. C., Romito, D., & Zanker, M. (2012, September). Linked open data to support content-based recommender systems.*Proceedings of the 8th International Conference on Semantic Systems* (pp. 1-8). ACM. doi:10.1145/2362499.2362501

Downes, S. (2014). Connectivism as Learning Theory. Half an Hour.

Drachsler, H., Hummel, H. G., & Koper, R. (2008). Personal recommender systems for learners in lifelong learning networks: The requirements, techniques and model. *International Journal of Learning Technology*, *3*(4), 404–423. doi:10.1504/IJLT.2008.019376

Fardinpour, A., Pedram, M. M., & Burkle, M. (2014). Intelligent Learning Management Systems: Definition, Features and Measurement of Intelligence. *International Journal of Distance Education Technologies*, *12*(4), 19–31. doi:10.4018/ijdet.2014100102

Ferguson, R., & Sharples, M. (2014). Innovative pedagogy at massive scale: teaching and learning in MOOCs. In *Open Learning and Teaching in Educational Communities* (pp. 98–111). Springer International Publishing. doi:10.1007/978-3-319-11200-8_8

Ganesh, S., Jayaraj, M., Kalyan, V., Murthy, S., & Aghila, G. (2004, April). Ontology-based web crawler. *Proceedings of the International Conference on Information Technology: Coding and Computing ITCC '04* (Vol. 2, pp. 337-341). IEEE. doi:10.1109/ITCC.2004.1286658

Gauch, S., Chaffee, J., & Pretschner, A. (2003). Ontology-based personalized search and browsing. *Web Intelligence and Agent Systems*, *1*(3-4), 219–234.

Goodwin, B., & Miller, K. (2013). Evidence on flipped classrooms is still coming in. *Educational Leadership*, *70*(6), 78–80.

Guarino, N., Masolo, C., & Vetere, G. (1999). Ontoseek: Content-based access to the web. *Intelligent Systems and Their Applications, IEEE*, *14*(3), 70–80. doi:10.1109/5254.769887

Hafidi, M., & Bensebaa, T. (2015). Architecture for an Adaptive and Intelligent Tutoring System that Considers the Learners Multiple Intelligences. *International Journal of Distance Education Technologies*, *13*(1), 1–21. doi:10.4018/ijdet.2015010101

Hanley, G. L. (2015). MOOCs, merlot, and open educational services. *MOOCs and Open Education Around the World*, 33.

Harasim, L. (2012). *Learning theory and online technologies*. Routledge.

Holotescu, C., Grosseck, G., Cretu, V., & Naaji, A. (2014, October). Integrating MOOCs in Blended Courses. *Proceedings of the International Scientific Conference eLearning and Software for Education* (*Vol. 4*, p. 243). "Carol I" National Defence University.

Iiyoshi, T., & Kumar, M. V. (2008). *Opening up education: The collective advancement of education through open technology, open content, and open knowledge*. Mit Press.

Jordan, K. (2015). MOOC completion rates, 2015. Retrieved from http://www.katyjordan.com.MOOCproject.html

Khosla, S., & Bhojane, V. (2013). Performing Web Log Analysis and Predicting Intelligent Navigation Behavior Based on Student Accessing Distance Education System. In Advances in Computing, Communication, and Control (pp. 70-81). Springer Berlin Heidelberg. doi:10.1007/978-3-642-36321-4_7

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: In search of the meaning of cognitive tools. *Instructional Science*, *35*(3), 207–256. doi:10.1007/s11251-006-9005-2

Kim, J. K., Kim, H. K., & Cho, Y. H. (2008). A user-oriented contents recommendation system in peer-to-peer architecture. *Expert Systems with Applications*, *34*(1), 300–312. doi:10.1016/j.eswa.2006.09.034

Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education*, *56*(3), 885–899. doi:10.1016/j.compedu.2010.11.001

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, *46*(5), 604–632. doi:10.1145/324133.324140

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *The International Review Of Research In Open And Distributed Learning*, *12*(7), 74–93. doi:10.19173/irrodl.v12i7.1041

Kumari, T., Gupta, A., & Dixit, A. (2014). Comparative Study of Page Rank and Weighted Page Rank Algorithm. *Proceedings of International Journal of Innovative Research in Computer and Communication Engineering*, *2*(2).

MacDonald, M. (2015). The Battle for Open by Martin Weller. *Journal Of Perspectives In Applied Academic Practice*, *3*(1). doi:10.14297/jpaap.v3i1.139

Malloy, T. E., & Hanley, G. L. (2001). MERLOT: A faculty-focused Web site of educational resources. *Behavior Research Methods, Instruments, & Computers*, *33*(2), 274–276. doi:10.3758/BF03195376 PMID:11447683

McGreal, R., Kinuthia, W., Marshall, S., & McNamara, T. (2013). *Perspectives on Open and Distance Learning: Open Educational Resources: Innovation*. Research and Practice.

Milligan, C., Littlejohn, A., & Margaryan, A. (2013). Patterns of engagement in connectivist MOOCs. *MERLOT Journal of Online Learning and Teaching*, 9(2).

Monge, S., Ovelar, R., & Azpeitia, I. (2008). Repository 2.0: Social dynamics to support community building in learning object repositories. *Interdisciplinary Journal of E-Learning and Learning Objects*, *4*(1), 191–204.

Ngambi, D., & Bozalek, V. (2015). Editorial: Massive open online courses (MOOCs): Disrupting teaching and learning practices in higher education. *British Journal of Educational Technology*, *46*(3), 451–454. doi:10.1111/bjet.12281

ODonnell, E., Lawless, S., Sharp, M., & Wade, V. P. (2015). A Review of Personalised E-Learning: Towards Supporting Learner Diversity. *International Journal of Distance Education Technologies*, *13*(1), 22–47. doi:10.4018/ijdet.2015010102

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: bringing order to the Web.

Paireekreng, W., & Wong, K. W. (2010, January). Mobile content personalisation using intelligent user profile approach. *Proceedings of the Third International Conference onKnowledge Discovery and Data Mining WKDD'10* (pp. 241-244). IEEE. doi:10.1109/WKDD.2010.119

Parr, C. (2013). MOOC completion rates below 7%. *Times Higher Education*, 9.

Passant, A., Samwald, M., Breslin, J., & Decker, S. (2009). Federating distributed social data to build an interlinked online information society.

Patil, A. S., & Pawar, B. V. (2012, March). Automated classification of web sites using Naive Bayesian algorithm. *Proceedings of the International MultiConference of Engineers and Computer Scientists* (Vol. 1).

Phillips, R., Preston, G., Roberts, P., Cumming-Potvin, W., Herrington, J., Maor, D., & Gosper, M. (2010). Using academic analytic tools to investigate studying behaviours in technology-supported learning environments.

Porcello, D., & Hsi, S. (2013). Crowdsourcing and curating online education resources. *Science*, *341*(6143), 240–241. doi:10.1126/science.1234722 PMID:23869007

Quercia, D., Askham, H., & Crowcroft, J. (2012, June). TweetLDA: supervised topic classification and link prediction in Twitter.*Proceedings of the 4th Annual ACM Web Science Conference* (pp. 247-250). ACM. doi:10.1145/2380718.2380750

Rai, L., & Chunrao, D. (2016). Influencing Factors of Success and Failure in MOOC and General Analysis of Learner Behavior. *International Journal of Information and Education Technology*, *6*(4), 262–268. doi:10.7763/IJIET.2016.V6.697

Rensing, C., de Freitas, S., Ley, T., & Muñoz-Merino, P. J. (Eds.). (2014, September 16-19). Open Learning and Teaching in Educational Communities. *Proceedings of the9th European Conference on Technology Enhanced Learning, EC-TEL 2014*, Graz, Austria, LNCS (*Vol. 8719*). Springer. doi:10.1007/978-3-319-11200-8

Salton, G., & Michael, J. (1983). Introduction to modern information retrieval.

Sathiyamurthy, K., & Geetha, T. V. (2012). Automatic Organization and Generation of Presentation Slides for E-Learning. *International Journal of Distance Education Technologies*, *10*(3), 35–52. doi:10.4018/jdet.2012070103

Schmidt, S., Scholl, P., Rensing, C., & Steinmetz, R. (2011). Cross-lingual recommendations in a resource-based learning scenario. In *Towards Ubiquitous Learning* (pp. 356–369). Springer Berlin Heidelberg. doi:10.1007/978-3-642-23985-4_28

Şenyuva, E., & Taşocak, G. (2014). Implementation of Web-Based Distance Education in Nursing Education in Turkey: A Sample Lesson in Patient Education. *International Journal of Distance Education Technologies*, *12*(3), 1–13. doi:10.4018/ijdet.2014070101

Sharples, M., Adams, A., Ferguson, R., Gaved, M., McAndrew, P., Rienties, B., ... & Whitelock, D. (2014). Innovating Pedagogy.

Siemens, G. (2005). Connectivism: A Learning Theory for the Digital Age. Retrieved from http://www.elearnspace.org.Articles/connectivism.htm

Smyth, R., Bossu, C., & Stagg, A. (2015). *Toward an Open Empowered Learning Model of pedagogy in higher education. Open learning and formal credentialing in higher education: Curriculum models and institutional policies*. Hershey, PA, USA: IGI Global.

Stacey, P. (2013). The Pedagogy Of MOOCs. The International Journal for Innovation and Quality in Learning.

Stern, M. K., & Woolf, B. P. (2000). Adaptive content in an online lecture system. In P. Brusilovsky, O. Stock, & C. Strapparava (Eds.), *Adaptive Hypermedia and Adptive Web-based systens* (pp. 225–238). Berlin: Springer-Verlag. doi:10.1007/3-540-44595-1_21

Tang, J., Wu, S., Sun, J., & Su, H. (2012, August). Cross-domain collaboration recommendation.*Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1285-1293). ACM. doi:10.1145/2339530.2339730

Toledo, R. Y., & Mota, Y. C. (2014). An e-Learning Collaborative Filtering Approach to Suggest Problems to Solve in Programming Online Judges. *International Journal of Distance Education Technologies*, *12*(2), 51–65. doi:10.4018/ijdet.2014040103

Town, C. (2007). The Cape Town open education declaration: Unlocking the promise of open educational resources.

Väljataga, T., Põldoja, H., & Laanpere, M. (2011). Open online courses: Responding to design challenges. Stanford University, H-STAR Institute, USA.

Van Harmelen, M. (2006, July). Personal Learning Environments. *Proceedings of ICALT* (Vol. 6).

Vassileva, J. (2008). Toward social learning environments. Learning Technologies. *IEEE Transactions on*, *1*(4), 199–214.

Wanaskar, U., Vij, S., & Mukhopadhyay, D. (2013). A Hybrid Web Recommendation System Based on the Improved Association Rule Mining Algorithm. arXiv preprint arXiv:1311.7204

Wang, J., Kraska, T., Franklin, M. J., & Feng, J. (2012). Crowder: Crowdsourcing entity resolution. *Proceedings of the VLDB Endowment*, *5*(11), 1483–1494. doi:10.14778/2350229.2350263 PMID:23355955

Wang, L. C., Meng, X. W., & Zhang, Y. J. (2012). Context-aware recommender systems. *Ruanjian Xuebao/ Journal of Software*, 23(1), 1-20.

Yair, Y. (2014). Open educational resources: reasons to be cheerful? *ACM Inroads*, 5(4), 37-38. http://doi.acm.org/10.1145/2684721.2684729

Zaïane, O. R., Li, J., & Hayward, R. (2006). Mission-based navigational behaviour modeling for web recommender systems. In *Advances in Web Mining and Web Usage Analysis* (pp. 37–55). Springer Berlin Heidelberg. doi:10.1007/11899402_3

Benneaser John is an experienced leader and architect in Software Engineering, who has worked in the Education Technologies building large scale Adaptive Learning Management systems. He is doing research work in the areas of Open Social Learning. Currently, Ben is heading the software engineering and platform architecture teams at AppNexus based in New York, USA.

J. Jayakumar is currently working as Associate Professor at Karunya University, Coimbatore, India.

V. Thavavel received her Master of Computer Applications in 1999 and Ph.D. in Computer Sciences in 2010, both from Madurai Kamaraj University, India. She is currently an Assistant Professor in the Department of Computer Sciences at the Prince Sultan University, Saudi Arabia, Riyadh. With 15 years of academic and research experiences, she has contributed more than 50 technical research papers in areas of medical image processing, optimization techniques and Big Data analytics. She has been distinguished guest speaker for IEEE and ACM society events. She is serving as a member on the review board of Elsevier and Interscience publications.

Muthukumar Arumugam has experience of 23 years of teaching computer science graduates. He also has 5 years of software industry experience. His areas of research includes Intelligent Tutoring Systems and Data Mining.

K. J. Poornaselvan has a PhD and is an Assistant Professor of Electrical and Electronics Engineering at Government College of Technology, Coimbatore, India.