Pan-Commonwealth Forum, 9-12 September 2019, Edinburgh, Scotland.

Directory of Open Educational Resources (DOER): A discovery service framework to provide structured access to OERs

Muthu Madhan[1] and Ricky Cheng[2]

[1]Azim Premji University, Bengaluru, India; madhan.m@azimpremjiuniversity.org
[2]Commonwealth of Learning, Vancouver, Canada; rcheng@col.org

Abstract

This paper discusses in detail the discoverability problem of OERs which is considered to be one of the major hurdles of user uptake. The paper also discusses the challenges in creating OER discovery services in an environment where non-interoperable OER repositories proliferate. It highlights some of the OER discovery initiatives at the community, national and international levels, and the case of Directory of Open Educational Resources (DOER), a structured discovery platform set up by Commonwealth of Learning (COL) is discussed in detail. DOER has brought most of the OERs available in repositories that are distributed across Commonwealth region into a structured metadata framework and it exposes the curated metadata to other service providers to enhance visibility. In perspective, DOER has demonstrated a frugal approach to creating open infrastructure for OER discovery by exploiting existing open standards and technologies. To solve the problems in OER discovery, authors suggest that OER advocacy should focus on strengthening discoverability features of institutional OER repositories by persuading them to follow open interoperability standards.

Introduction

Only a few ideas in the recent past have had as great an influence as Open Educational Resources (OERs) have on education. In the keynote address at the MERLOT Annual Conference 2005, Sir John Daniel et. al. said "the combination of rapidly increasing connectivity across the world and the emergence of the concept of reusable learning objects - specifically open educational resources - have tremendously hopeful implications for humankind."[1]

Significant spending by philanthropic foundations and government agencies on development of OERs and OER repositories, and the determined efforts of institutions across the world have helped a lot in making web-wide availability of OERs in high volumes. Despite these developments, challenges and barriers exist in connecting OERs and users. Discovery of OERs has been one of the major barriers of user uptake. A recent report which examines aspects of adoption of OERs in 21 countries in South America, Sub-Saharan Africa and South and Southeast Asia says locating OERs to use and build on is often a time-consuming process for educators.[2] Structured resource description and interoperability for seamless data exchange are essential features of open repositories and unfortunately, many repositories that are populated with OERs lack these features.

In 2014, *IEEE Transactions on Engineering* brought out a special issue on Open Educational Resources in Engineering Education. Editors of the special issue 'from their over 6 years of experience leading a research group in semantic Web technologies applied to Open Education' offered some recommendations with a view to making OER data easily accessible by both human and machine. A key requirement was, in their opinion, "to improve the metadata interoperability between various collections of open material, so as to facilitate discoverability".[3] The 'Open Educational Resources: Global Report 2017' – a report based on a stakeholder survey has emphasized adding discoverability features would be important for stakeholders successfully mainstreaming OER in teaching and learning.[4] In spite of these recommendations, interoperability among OER repositories remains an issue and the problem is largely technical. GÉANT, the pan-European data network for the research and education community, has stated that wider take-up and use of Open Education Resources largely depends on how technology can support in offering OER facilities a financially sustainable manner.[5]

What are the problems in OER discovery?

In general OER repositories around the world are functioning in silos and most of them are not optimised for discoverability. The online repositories that are populated with OERs have not focused on adopting open metadata standards to describe OERs, and they have not adopted mechanisms for seamless information exchange between systems. Therefore, the individual OER repositories are not interoperable and by design they could not form part of a web-wide network. For example, take some OER repositories in India which are populated with valuable contents viz. NPTEL <https://nptel.ac.in/> Swayam < https://swayam.gov.in/ >, NROER < https://nroer.gov.in/home/repository >, Egyankosh < http://egyankosh.ac.in/ > and Teachers of India < http://teachersofindia.org >. Not all these repositories are interoperable. Therefore creating a discovery platform by seamlessly harvesting metadata from these repositories becomes a challenge. This problem is not specific to Indian OER repositories.

Aggregating metadata from non-interoperable repositories would require manual intervention and proliferation of such OER repositories makes creation of discovery platforms highly resource intensive. However, there are efforts at various levels to aggregate OERs from individual repositories and create discovery services. Some of such initiatives are highlighted here.

As early as 2007, the Institute for the Study of Knowledge Management in Education (ISKME) launched OER commons < https://www.oercommons.org/ >, a discovery infrastructure for curriculum experts and instructors at all levels to identify high-quality OER and collaborate around their adaption and evaluation. OER commons has aggregated more than 48,000 OER objects (data as seen on 13 August 2019) from 350 OER providers, and most of them are from the USA and a few other developed countries. In addition to aggregating metadata, OER commons has offered features such as social bookmarking, tagging, rating, and reviewing to engage users with OERs available in the platform.[6] Hewlett Foundation, one of the supporters of ISKME initiatives, has funded USD 7 million to ISKME during 2003-2019, of which approximately USD 1.7 million has been spent for the development of OER commons platform.[6]

Also, Hewlett Foundation has funded USD 898,000 to North Rhine-Westphalian Library Service Center (hbz)[6] for the development of OER Worldmap < https://oerworldmap.org/ > which aims to provide a global map of OER organisations, projects, people and services. The platform allows users identifying of OER collections from its qualified lists of repositories and it facilitates social networking. And, the platform maintains a registry of OER policies developed across the world.

In 2015, the National Mission on Education through Information and Communication Technology (NMEICT), a Government of India (GoI) project, has created a learning resources discovery platform called National Digital Library of India (NDLI). < https://ndl.iitkgp.ac.in/ > This platform is trying to aggregate contents available in digital repositories from India and elsewhere. GoI will be spending about USD 7.5 million (or INR 50 crore) on NDLI to strengthen its content and technology, and to develop a sustainable business model to run the platform.[7] While NDLI aggregates contents easily and seamlessly from interoperable institutional digital repositories, it requires huge amount of manual intervention to harvest, curate and index metadata from non-interoperable repositories and to include offline contents. After four years of experimentation, members of the NDLI project implementation team have emphasized the need for a credible and open mechanism for content sharing and collaboration.[8]

In 2016, GÉANT launched a pilot OER metadata harvester called eduOER < https://oer.geant.org/ > which aims to harvest metadata from repositories across Europe using the *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) and provide a structured discovery interface. eduOER picks up the results of the OER pilots such as Terrena[9], which was accomplished at the cost of Euro 39,000.[10] The main aim of eduOER is to create an open harvester infrastructure which will harvest and index metadata from European OER repositories and expose them to all possible would be users across the world. One of the unquantifiable benefits expected from the service is 'increased number of lifelong learners in Europe and globally (The University of Life)'.[11] eduOER has harvested more than 65,000 OERs available (in audio and video formats) in nine languages including Arabic and Hebrew and its harvest goes beyond Europe to include some US academic repositories. GÉANT was seeking for way to secure further funding for eduOER.[9]

Directory of Open Educational Resources (DOER)

Background

OER advocacy has been one of the main focus areas of COL since 2005, and it has produced more than 250 papers on various aspects of developments in open educational systems, services and practices (data as seen in http://oasis.col.org). Particularly, COL was tracking and reporting OER developments in the Commonwealth region through periodical information bulletins.

In 2011, with a view to increasing web-wide visibility of contents available in OER repositories set up by universities, research and educational institutions in the Commonwealth region, COL wanted to develop a Directory of Open Educational Resources (DOER). <http://doer.col.org/> Main objective of the Directory was to harvest and index metadata (with links to original source) from OER repositories that are interoperable and that are not, and create a discovery interface with multiple choices of access points for users. More importantly, the COL team wanted to achieve it through frugal means, and at the same time with no compromise on the requirements for a scalable open scholarly infrastructure.

Technology

COL team learnt that OAI-PMH had already emerged as a practical solution for interoperability among digital repositories and many open source repository software were inherently OAI-PMH compliant. Compliance with the protocol can make a repository interoperable and it helps the repository move metadata records to search engines seamlessly.

After careful analysis of several open source platforms, COL team decided to use DSpace < https://duraspace.org/dspace/ > – open source digital repository software – to build the directory. DSpace is a flexible and customizable platform comes with configurable metadata schemas, workflows and, browse and faceted searching functionalities. It complies with standard protocols and best practices for access, ingest, and export data. DSpace platform is available in more than 20 languages. One can store any type of file in DSpace. And, it has had a strong community involvement. Many academic and non-profit organizations across the world have used DSpace to build open digital repositories. The Directory of Open Access Repositories (Open DOAR) maintained by the Joint Information and Systems Committee (JISC), UK has listed more than 4,000 OAI-PMH compliant repositories, of these 43% are built on DSpace.[12] These factors went in favour of DSpace.

DOER is evolving in terms of content and technology ever since it was setup in 2011. DSpace was extensively customized, and the in-built metadata schema and input forms were enhanced to suit the needs of DOER. Initially, DOER was built using DSpace version 1.8.x and it was upgraded to version 5.5 in 2017 to exploit the features in the new version of the platform. After the upgrade DOER was re-launched in 2018.

Metadata acquisition

Administrators of the Directory have identified about 7,200 OERs published during 1995-2018 which are located in more than 200 repositories from different regions of Commonwealth. < http://doer.col.org/browse?type=title > These repositories are not interoperable, therefore, it required human intervention to curate metadata of each OER gathered from these repositories. DOER development team had to identify and fill missing metadata elements wherever possible. Persistent identifier for each record was ensured.

A record in DOER captures metadata of an open course which may contain multiples module and lectures. For example, one NPTEL course (say Engineering Physics 1) has 40 video lectures under eight modules. < https://nptel.ac.in/courses/122107035/>. DOER maintains only one metadata record for this course < http://doer.col.org/handle/123456789/5475>, and it directs users to the original NPTEL page where one can get links to all the 40 lectures. DOER has not created metadata for the individual modules or lectures. The 974 NPTEL records that DOER has indexed would provide links to approximately 35,000 lectures.

DOER has organised these resources in more than 20 OER collections under four broad communities viz. Higher education (4267 records), Open schooling (1,054 records), Teacher education (860 records), and Technical & Vocational Skills development (1,026 records).

Harvesting from interoperable repositories

Late versions of DSpace come with a harvester plugin which enables metadata crosswalks between DSpace repositories. Some institutional OER repositories in the Commonwealth region have been built using DSpace. DOER can seamlessly harvest metadata from such repositories and ingest into it. For example the Indira Gandhi National Open University, India has set up a DSpace repository called eGyanKosh < http://egyankosh.ac.in > and opened access to self-instructional print material of its 2565 courses and 2389 video programmes.[13] The DOER team has successfully tested harvesting metadata of more than 34,000 records from eGyanKosh, and all these records can be moved to the production server of DOER. Even from the OAsis, the institutional repository of COL, metadata can be harvested and indexed seamlessly. DOER can incrementally and periodically harvest metadata, so, updates in the source repository is reflected in DOER as soon as possible.

Also, DOER considers expanding its scope by harvesting metadata from important OER repositories located in other regions. DOER team has established metadata exchange agreement with SkillsCommons < www.skillscommons.org/ > - a DSpace based digital library of workforce training materials, set up by the US Department of Labor. DOER will soon harvest the 16,500 records available in SkillsCommons.

Use of DOER

DOER has a powerful search engine that indexes all metadata elements of each OER record and text files attached to it. One can search the records in the directory using simple and advanced search features and, filter the results by facets such as creators and publishers.

DOER captures statistics related to the number of searches performed (using the DOER interface), the top 10 most often used search terms, the top 10 countries and cities from where the search or view originates, and the top 10 most viewed records. One can get these statistics at the level of communities, collections and individual records. When DOER software is upgraded more statistical features may be expected.

In all, more than 74,000 searches (data as seen on 18 June 2019) have been performed using the DOER interface. 'Teaching and learning methods', 'Teacher education and training', 'Curriculum management', 'Primary education', and 'Mathematics' are the most often used search terms. Frequently used search terms in DOER distributed by community is given in Table 1. These searchers led to about 2000 page views during June 2018 – May 2019. The views originate from many countries including the United States, India, Canada, Russia, France, South Africa, China, United Kingdom, Uganda, and Namibia. Number of views from different countries distributed by DOER community is given in table 2. Saint Petersburg, Ashburn, Vancouver, Mountain View, Bangalore, Noida, Pretoria, Beijing, Windhoek, and Calgary are the cities from where the views have originated. (Table 3).

Future work

DOER is an in-house project of COL and a small team is involved in populating and maintaining the platform. The team will continue to enhance features of DOER and focus on challenges in harvesting, curating and de-duplicating metadata harvested from different types of repositories. And, the team will explore ways to use technologies cleverly and effectively address the issues in the present OER ecosystem. DOER team will make concerted efforts to conduct outreach programmes for OER producers and will grab opportunities for collaboration with other service providers to enhance the OER discovery systems at the global level.

Discussion

DOER has shown how existing technologies can be exploited in creating an open discovery infrastructure, and it has also demonstrated a frugal approach for a structured OER discovery. DOER has brought most of the OERs available in distributed and non-interoperable repositories located in the Commonwealth region into a structured metadata framework. It exposes the curated metadata to other service providers which enhances visibility of contents indexed in DOER. < https://www.base-search.net/Search/Results?q=dccoll%3Aftdoer&refid=dclink > With minimal communication efforts and virtually no special programmes to popularize DOER, it has registered a reasonable usage.

The value of interoperable repositories is increasingly acknowledged by the advocates of open access (OA) movement which aims to provide unrestricted online access to research outputs (such as journal articles) produced from different parts of the world. Major research funding agencies across the world have mandated researchers they fund to archive their research outputs in interoperable repositories. Bielefeld Academic Search Engine (BASE) which is maintained by Bielefeld University Library is harvesting metadata from 6,000 repositories and it has harvested about 140 million records so far and indexed them with enhancements.

The global OER advocacy has focused for long and too much on open contents. What is important is, as the Open Educational Resources: Global Report 2017 has emphasised, to focus on strengthening discoverability features of institutional OER repositories by persuading them to follow open interoperability standards.[4]

Reference

1. Daniel, J.,West, P.,Antoni D, S. & Uvalić-Trumbić, S. The Long-Awaited Breakthrough? http://hdl.handle.net/11599/1541 (visited on 26, June 2019)
2. Hodgkinson-Williams, C. & Arinto, P. B. (2017). Adoption and impact of OER in the Global South. Cape Town & Ottawa: African Minds, International Development Research Centre & Research on Open Educational Resources. doi: 10.5281/zenodo.1005330
3. Tover, E & Piedra N, (2014), Open Educational Resources in Engineering Education: Various Perspectives Opening the Education of Engineers, *IEEE Transactions on Education*, 57 (4), 213-219 doi: 10.1109/TE.2014.2359257
4. Commonwealth of Learning (2017). Open Educational Resources: Global Report 2017. Burnaby: Commonwealth of Learning http://hdl.handle.net/11599/2788 (visited on 26, June 2019)
5. Geant, eduOER: The Open Educational Resource Hub & Portal service of GEANT https://oer.geant.org/about- eduoer/ (visited on 25 June 2019)
6. Hewlett Foundation, Grants database, https://hewlett.org/grants/?keyword=ISKME&sort=date&current_page=2 (visited on 25 June 2019)
7. Verma, P. (2018) India's digital library plans to collaborate with foreign peers, *Economic Times* https://economictimes.indiatimes.com/industry/services/education/indias-digital-library-plans-to-collaborate-with-foreign-peers/articleshow/63371973.cms?from=mdr (visited on 25 June 2019)
8. Das, PP, et. al. (2016) National Digital Library: A Platform for Paradigm Shift in Education and Research in India, *Science and Culture*, 4-11
9. EduOER Service development history. https://oer.geant.org/history/ (visited on 25 June 2019)
10. GEANT Open Educational Resource Portal pilot https://wiki.geant.org/display/tfmedia/Open+Educational+Resource+Portal+pilot#OpenEducationalResourcePortalpilot-del (visited on 25 June 2019)
11. Dias, N. et. al. (2016) Open Educational Resource (eduOER) Business Case Development, https://oer.geant.org/wp-content/uploads/2016/04/D12.1_GEANT-eduOER-Business-Case-Development.pdf (visited on 25 June 2019)
12. JISC, OpenDOAR statistics http://v2.sherpa.ac.uk/view/repository_visualisations/1.html (visited on 25 June 2019)
13. IGNOU, Reactivation of IGNOU's eGyanKosh http://ignou.ac.in/userfiles/eGyankosh-reactivation.pdf (visited on 25 June 2019)

Table 1. Frequently used search terms (top 7) in DOER distributed by community from 1 June 2018 – 30 May 2019

Higher Education		Open Schooling	
Search terms	**No. of searches**	**Search terms**	**No. of searches**
Mathematics	649	Mathematics	655
Computer science & software engineering	551	Geometry	501
Mechanics & thermodynamics	542	Chemistry	468
History	520	Biology	461
Mechanical engineering	512	English	437
Physics	505	Arithmetic	429
Open source software	443	Physics	393
Teacher education		TVSD	
Search terms	**No. of searches**	**Search terms**	**No. of searches**
Teaching and learning methods	1132	Indian Institute of Technology	346
Teacher education and training	930	Open source software	322
Curriculum management	886	Job search	308
Primary education	736	Computer literacy	307
Mathematics	659	Study skills	296
School & college administration	536	Building & construction	283
Secondary education	528	Programming language	260

Table 2. No. of views from different countries distributed by DOER community (1 June 2018 – 30 May 2019)

Country	Higher education	Open schooling	Teacher education	TVSD
United States	126	101	118	122
India	80	40	30	33
Canada	77	20	28	64
Russia	68	19	20	23
France	42	41	42	46
South Africa	41	8	14	0
China	27	17	21	13
United Kingdom	22	11	17	15
Uganda	14	-	-	-
Namibia	10	-	-	-
Phillipines	-	7	-	-
Guyana	-	4	-	-
Australia	-	-	5	20
Germany	-	-	5	7
Hong Kong	-	-	-	21

Total views from top 10 countries	507	268	300	364
All views	625	325	379	440

Table 3. No. of views from different countries distributed by DOER community (1 June 2018 – 30 May 2019)

City	Higher education	Open schooling	Teacher education	TVSD
Saint Petersburg	64	19	19	22
Ashburn	49	48	56	45
Vancouver	44	6	10	40
Mountain View	20	18	18	21
Bangalore	19	9	7	-
Noida	18	-	8	17
Pretoria	18	-	0	-
Beijing	14	5	10	-
Windhoek	10	-	-	-
Calgary	7	-	-	-
Ann Arbor	-	8	6	7
Freemont	-	5	13	-
Guargoan	0	5	0	0
Windsor	0	0	6	0
Sydney	0	0	0	17
Stoney Creek	0	0	0	12
Kirkkonummi	0	0	0	7
Total	263	123	153	60