

Improving the Creation, Curation and Discovery of Open Educational Resources for Work-Integrated and Lifelong Learning

William Jobe¹, Stefan Nilsson¹, Patrik Bernhardsson¹ and Anders Bindslev Rask²

¹ University West/Informatics, Trollhättan, Sweden

² University College Nordjylland, Aalborg, Denmark

Abstract— This paper concerns open education (OER) resources, and more specifically the process of publishing OER material to make it easier to find on the web. The overarching goals are to improve the creation and discovery of OER material for workplace learning as well as strive towards UNESCO’s Sustainable Development Goal 4, which aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. As we see it, there are two key challenges. The first challenge regards the content creation and curation side; i.e. the process of making OER material available to the general public, how to maintain the resource and how to get feedback on it. The second challenge regards discovery i.e. for OER consumers to easily find OER material and discern its usefulness in their specific learning context.

We are in the process of developing a web based system that aids content creators in publishing and maintaining their OER material. The system helps the content creator by generating semantic metadata about the resource based on the Dublin Core system and creates a package of the material itself along with said metadata in a search engine optimized HTML5 file, ready for publishing on a web server or a video streaming site like YouTube, or other media sharing sites.

This approach primarily utilizes common search engines to find the material, thus making the OER material decentralized and not necessarily specific to a specific OER repository. The metadata created by the system does, however, enable developers to create systems specialized in collecting and aggregating OER material, further enhancing the capability of the decentralized ecosystem of OER resources to be used by learning management systems. The key contribution is a suggestion as to how to automatically generate and use an OERID so that all resources can be discovered, curated, and reused.

Index Terms—Discoverability, OER, Dublin Core, HTML5

I. INTRODUCTION

The term OER was first coined at the UNESCO forum in 2002 to refer to the need for a universal educational resource available for all of humanity [1]. Since the inception of this concept, the term OER has come to mean a resource for “teaching, learning and research materials in any medium that reside in the public domain and have been released under an open license that permits access, use, repurposing, reuse, and redistribution by others with no or limited restrictions” [2]. This definition provides the key ingredients of OER, which are that they are for learning and teaching, open and reusable and exist in a variety of

mediums. Over the last fifteen years, a variety of private, public and open stakeholders have initiated numerous OER initiatives in the form of portals, databases and repositories [3].

The growth and use of discrete OER repositories has created several problems. Many overlapping and competing standards and educational metadata schemas have been used and proposed overtime, including IEEE LTSC LOM (<http://ltsc.ieee.org/wg12/par1484-12-1.html>) (Learning Object Metadata), one of the widest adopted, IMS (www.imsglobal.org/metadata/), Ariadne, ISO/IEC MLR - ISO 19788 (www.iso.org/iso/), Metadata for Learning Resources (MLR) and Dublin Core [4]. Furthermore, the web consists of a multitude of platforms that are accessed from a variety of devices and several metadata schemas are used by OER repositories and these repositories offer very heterogeneous datasets, differing with respect to schema, exploited vocabularies, and interface mechanisms [4].

The need to address the issues of heterogeneous standards and schemas is well-known and “accessibility is absolutely vital for a project to produce truly ‘open’ educational resources” [5]. Previous research presents a few specific issues with OER. Four major challenges for OER according to [4] are integrating distributed data from heterogeneous educational repositories, dealing with continuous change, how to structure texts and evolve taxonomies, and how to compile metadata to facilitate web discovery at scale. According to [6] another current issue with OER is that there is no mobile distribution platform. [7] further mentions the size of the search pool and the ability to filter by resource properties as issues that must be addressed by OER search tools. Finally, [4] point out that web searches of educational resources face a heterogeneous landscape and individual repositories and metadata stores largely use XML and relational databases.

OER is used in several different contexts. One such context is the area of Work-Integrated Learning (WIL). [8] propose how the OER movement can become a key foundation for curating WIL practices which are mindful of disciplinary heritage, local institutional contexts and fostering a capability approach. “The notion of curating work-integrated learning via the OER movement is proposed as a way of not just preserving the past, but also as a way of innovating present and future practices as well” [8].

II. BACKGROUND AND OER RESEARCH OVERVIEW

Some previous research focuses on the use of different metadata systems to tag OER and addresses issues of

discovery and curation. The LOM-based search approach is an opportunity for teachers and students to teach and learn together in a collaborative learning environment. The selected fields from the LOM standard enable access to educational resources in different web-based search systems in a formal and effective manner [9]. Linked Data technologies enable the development of the next generation of OER, allowing the separation of semantics from syntax, the improvement of discoverability and access, and the use of common vocabularies, and additionally, the proposed architecture provides an opportunity to merge data distributed across different libraries [10]. The CROERA (Cross-Repository Open Educational Resources Aggregation) system is an aggregator of educational resource repositories that enables access to resources independently of the taxonomy utilized to classify the content in each repository. This enables users accustomed to the descriptions and classification strategy of a particular repository to discover resources in other repositories without needing to switch to another taxonomy or classification scheme. This platform was designed to integrate any existing repository [11].

While the metadata schemes evolve, issues have been raised as to the quality of the metadata as entered by users (see for example [12], [13]. [13] also note that this is an issue that has remained true for an extended time, citing previous research findings reporting that only a fraction of the available fields from various metadata schemes was used to identify and describe an OER.

[5] found that search engine optimization (SEO), meshing existing OER resources with social web functions and recommendation systems are key avenues to explore for OER discoverability. An “appropriate OER” must be understood as a document that at least allows re-purposers to modify contents on both the legal (license) and the technical (format) level [14]. [14] also propose that the possibility to change an OER resource should be mandatory in the metadata, that an OER object comes in both the original, unchangeable format and a reusable format and that all repurposed resources are made public to ensure continuity. Regarding perceptions and attitudes towards OER, [15] found that key areas to address were the need for more information and a need for central repositories that are easily discoverable. Concerning mobile authoring tools for OER creation, [16] presented and tested a mobile application MAAIMS (Mobile Authentic Authoring in IMS) that enabled content creation that was shareable via a repository and reusable in multiple contexts. Finally, [17] concluded that web 2.0 tools improve OER searches by using their semantic meaning. The social and semantic components of the OER production cycle reduces the time needed for development of resources, and including social and semantic components in OER production is without a doubt an opportunity to take advantage of the benefits of web-related trends that provide pertinence, relevance and meaning to the resources.

A. Creation, Curation and Discovery

The complexity of the issues for the dissemination and adoption of OER can be reduced to three key issues: creation, curation and discovery. In an OER context creation is the process of creating, tagging, licensing, packaging and sharing an OER object, curation is the process of managing, reusing and adapting an OER resource, and discovery is the process of easily finding an

appropriate OER object. [6] describes that there is a lack of curation in the digital content ecosystem. [7] explains that publishers as well as educators consider discovery as a barrier to OER adoption. The same conclusion is presented by [18], listing “No comprehensive catalog” and “Too hard to find what I need” as the two top reasons deterring educators from the use of OER. Furthermore, [7] states that the three key ingredients for an OER discovery tool are curation, indexing and metadata exchange.

One common solution to address OER creation, curation and discovery is to use search engines in combination with existing tools and technologies. The single most important function for metadata is to facilitate resource discovery which in a practical OER context means facilitating a web search [5]. As [19] explain, the use of search engines as a primary tool for discovering academic content should be discussed and addressed, when considering how to increase access to OER. An improved OER search tool should not require the adoption of specific technologies and should leverage the behavior of users, as well as help them refine and filter searches based on metadata as provided by the creator or a curator [7]. Further, a OER discovery tool should provide results that are relevant, usable and from multiple sources [7]. [20] proposes hosting OER in at least two locations to overcome the vulnerability of discrete web repositories, and by using SEO techniques and social media channels, OER can quickly become widely dispersed and discoverable. The ultimate aim is to unify the fragmented landscape for OER on the web [4], [21].

III. RESEARCH AIM

The first challenge regards creation and curation; i.e. the process of making OER material available to the general public and maintaining the resource. The second challenge is discovery; i.e. allow OER consumers to easily find OER material and discern its usefulness in a specific learning context. The aim of this research is to address creation, curation and discovery issues with OER by proposing a simple online tool that automates the process of creating OER objects.

The following sections describe the online tool as a proof of concept to address these longstanding issues with OER.

IV. RESULTS

Much in line with [20], [4] and [21], we propose an OER tool to address the identified issues of creation, curation and discovery by aiding content creators to automatically tag their OER resources relevantly and allowing the creators to host their OER resources on web hosting services or social media services. Research and suggestions from e.g. [5], [19] and [7] support the idea of sharing an OER resource on existing, open platforms that are easily accessible through search engines and not just specific OER repositories. In order to do so, we suggest an automatically generated OERID as a unique identifier for all OER objects. This OERID can then be used to track curation and simplify the discovery process. The focus on and use of an automatically generated unique OERID further addresses the problem with curation. The OERID will allow for any OER object to be tracked, repurposed and reused throughout the life of the resource and any changes can be logged and tracked as well. Finally, the combination of simplified, automated creation, tagging and publishing of a resource with an OERID directly addresses the three issues of creation, curation, and discovery. Table 1 summarizes

how the proposed prototype addresses the aforementioned issues.

A. *Prototype description*

The prototype presented here is a “proof of concept”, developed with the aim of utilizing it as a tool for

TABLE I.
OER ISSUES AND PROPOSED SOLUTIONS

Issue	Proposed solution from OER tool/prototype
Creation	Simplify and streamline creation and tagging process Automatically generate a unique OERID Allow for publication of the resource on open, social platforms
Curation	The use of a unique OERID allows for a resource to be tracked and thus curated, when reused or repurposed
Discovery	By using a unique OERID and open platforms accessible from common search engines, resources are easier to discover and link to

discovering the feasibility of addressing the aforementioned issues of creation, curation and discovery of OERs using contemporary web technologies. Key considerations for developing the prototype were thus to simplify the process of tagging the OER for the content creator and create metadata about the object in order to aid in the process of publishing a findable, well described OER resource.

Using standard web components, the prototype is driven on the backend by Node.js together with several open source JavaScript components allowing for the extraction of metadata about the OER object to automate the tagging process as much as possible. A few open web technologies are used on the frontend, predominantly Angular.js providing the user interface and JavaScript components allowing for the upload of files.

B. *Describing the OER*

To mark up the OER material, the content creator has to fill in a number of metadata fields describing the material. As previously stated, several metadata schemes for marking up the OER metadata exists, using more or less extensive and exhaustive sets of fields for the author to fill in. Although extensive, previous research has reported that the quality of the metadata on existing OERs within repositories is low, suggesting that there is a need to aid the content creator as much as possible in the process of describing the material.

Logging in using the content creator’s email address, the user is presented with the opportunity to upload the OER material (see Fig. 1). The upload feature is mainly used to extract the metadata from the material. At present, the prototype supports the most common image types, video files (in the form of mp4) and documents in the form of PDFs.

The screenshot shows the OER-ID user interface. At the top, it says 'Welcome to OER-ID!' with navigation links for Home, All, and Upload. A large box contains the instruction 'Drop files to upload (or click)'. Below this is a form for 'Your e-mail address' with the example 'jane.doe@anyuniversity.org' and a note 'We will never share your e-mail with anyone'. The main section is titled 'Dublin Core Metadata Element Set' and contains several input fields: Title, Subject, Description (with a placeholder 'Please describe the resource'), Date (set to 'Fri May 25 2018 18:06:06 GMT-'), Language (set to 'English'), Creator (set to 'Jane Doe'), Contributor (set to 'John Doe'), Publisher, Coverage, Rights (set to 'CC0'), Format, Type, Source, and Relation. At the bottom, there is a field for 'The OERID:' and two buttons: 'Reset all fields' and 'Download distributable package'.

Figure 1. The OERID user interface

After the file has been uploaded, the user can fill in the fields from the Dublin Core Metadata Element Set standard [22]. Depending on the amount of metadata that has been extracted from the file at upload, some of the fields are prefilled. The user can adjust these fields in case they are inaccurate. Table 2 outlines the fields according to the Dublin Core Metadata Element Set, and demarks fields that can a) be automatically filled in, or b) be suggested by extracting data from the web server, the user’s browser or the resource’s metadata. Further, the prototype can be personalized in such way that the user can add common contributors, coverage, publisher, and subject to be reused.

TABLE II.
OER ISSUES AND PROPOSED SOLUTIONS

Term	Short Description	Automated
Contributor	Name of person contribution to the resource	Suggested
Coverage	Names of places and dates the resource covers	Suggested
Creator	Name of the creator of the resource	Suggested
Date	A date of the resource	Yes
Description	A text describing the resource	No
Format	The file format or distribution form. Could also be dimensions of an image or duration of a video.	Yes
Identifier	A unique reference to the resource	Yes (OERID)
Language	The language associated with the resource	Suggested
Publisher	Name of organization or person publishing the resource	Suggested
Relation	ID of a related resource, could be an OERID	No
Rights	Property rights associated with the resource	Suggested
Source	Identity of a resource from which this resource is derived, could be an OERID	No
Subject	The overarching topic of the resource using keywords	Suggested
Title	The formal name of the resource	No
Type	Type of resource, such as "image" or "text"	Yes

An OERID is automatically generated for the file, tying together the user with a particular OER resource. The unique OERID is an alphanumeric string generated by the application as a MD5 hash using the user's e-mail address, the filename and file size of the resource, the type of the resource, together with a timestamp from the server at the time of upload. Finally, the alphanumeric string is prepended with the string "OERID". As such, the OERID can be used to validate the origin of the material. In extension, the OERID can also be used as a resource identifier by other services such as specialized OER search engines as well as for example curation and validation services.

C. Distribution of the OER

When the necessary fields are completed, the user submits the form. The system then generates a package of files for the user to download. The package contains an HTML5 document using semantic tags containing the metadata of the OER, as well as a link to the OER object itself. Creating a semantic HTML5 document helps with the SEO as well as the presentation of the resource itself. Using the files in the package, the content creator is able to distribute the material to a web location of their choice. In

the case that the OER is a video, the content creator can further use any video sharing service such as YouTube or Vimeo, and provide the unique OERID in the description.

D. Curation of OER

Using the OERID, collections of open educational resources can be curated by web services. Utilizing the OERID both as an identifier of a resource as well as using it as an identifier of a related resource and a resource from which it is a derivative of, one can track a single OER and how it relates to other OERs.

V. DISCUSSION

This paper has explored current issues relating to the use of OER, and identified three areas of interest: the creation, curation and discovery of OER material. A key discussion question is how to alleviate these issues. As [20] states "We will need to consider whether we wish to continue tracking the fate of our content, or whether those involved in global open education will have to accept that OER will fledge and leave their project nests as part of the natural cycle of events." Our proposed prototype eases creation and uses a unique OERID that offers a solution for curation and discovery. The prototype encourages OER content creators to share their OER resources on a variety of platforms, thus freeing the material from closed repositories. However, our suggested solution still allows for creators to track and make the material easily accessible from a general search engine by using an OERID together with a semantic HTML in a HTML5 format. We view this solution as the best of both worlds, i.e. OER material can be shared anywhere and easily found without losing the ability to track and curate the material.

Another issue of discussion is the use of OER resources in different contexts. There are several contexts such as e.g. work-integrated learning, lifelong learning (non-formal education), formal education and open education for fulfilling UNESCO's SDG goals, where research highlights the need for better utilization of OER resources and the possible benefits of increased OER use. Our proposed prototype has the potential to make headway in these contexts by increasing the discoverability and availability of OER resources. Furthermore, the ability to track, curate and repurpose OER resources using the OERID can allow for the development of informal communities in these contexts that can lead to increased quantity and quality of OER resources.

The concept of using a unique ID, the proposed OERID, for tracking all OER resources has both advantages, disadvantages and technical issues. A key advantage of the OERID is that it provides a way to uniquely identify and track an OER resource for the life of the resource. This is advantageous as it provides improved discoverability and curation. Another advantage is that the ability to track changes to an original resource provides transparency to any changes and repurposing. We also see the possibility for content creators to trace an OER to see where and in which contexts a specific resource they created is being used. This can also be used as a validation tool for teachers and educators looking for OER material, seeing where and by whom a resource has previously been used.

However, a possible disadvantage of the OERID is that it can decrease anonymity of a resource that can be desirable in certain contexts. Furthermore, there are some technical issues regarding format and "uniqueness" of the

OERID. Our prototype is a proof of concept and the algorithm used to create the OERID should be analyzed and discussed in more detail. We deem our proposed algorithm that automatically generates the OERID to be unique enough, but will it scale if use increases drastically? Would it be better to strictly create the OERID from exif data from a resource and randomly generated string without requiring user data (email in our algorithm)?

While the prototype automates certain aspects of describing an OER according to the Dublin Core Metadata Element Set standard [22], the results vary depending on the type and format of the resource. As the automation of describing the OER lowers, we expect that the quality of metadata follows the same pattern. One can argue that since the idea of this project is to allow for OERs to be publicly available on standard web services and trackable by the OERID, it could provide a breeding ground for applications and services further aiding the creation, curation and discoverability of OERs. One such application could be what [12] term collaborative metadata creation, where the content creators can gain assistance from metadata specialists as well as users in the OER community in improving the metadata quality.

VI. CONCLUDING REMARKS

The aim of this paper has been to explore how, by the use of technology, we can alleviate the hurdles of creating, curating and discovering OER material. The prototype developed aids the content creator in the creation stage by automating the process of filling in the fields from the Dublin Core Metadata Element Set standard [22]. Further, it creates a unique identifier for each OER resource, dubbed the OERID, assisting content owners in the curation of the resource they have provided the OER community. When it comes to the discovery stage, the application creates a standardized HTML5 file that accompanies the OER resource. Utilizing semantic tagging of the provided DCME fields helps search providers such as Google and Bing to find and rank the OER material. The OERID can also be used in the description field of video streaming sites such as YouTube or Vimeo to identify the resource.

REFERENCES

- [1] UNESCO. (2002). Forum on the impact of open courseware for higher education in developing countries (pp. 1–30). Retrieved from http://portal.unesco.org/ci/en/ev.php-URL_ID=2492&URL_DO=DO_TOPIC&URL_SECTION=201.html
- [2] Atkins, D. E., Brown, J. S., & Hammond, A. L. (2007). A review of the Open Educational Resources (OER) movement: Achievements, challenges, and new opportunities, 1–84.
- [3] Wang, T., & Towey, D. (2017). Open educational resource (OER) adoption in higher education: Challenges and strategies (pp. 317–319). IEEE. <https://doi.org/10.1109/TALE.2017.8252355>
- [4] Dietze, S., Sanchez-Alonso, S., Ebner, H., Qing Yu, H., Giordano, D., Marenzi, I., & Pereira Nunes, B. (2013). Interlinking educational resources and the web of data: A survey of challenges and approaches. *Program*, 47(1), 60–91. <https://doi.org/10.1108/00330331211296312>
- [5] Thomas, A., Campbell, L. M., Barker, P., & Hawksey, M. (Eds.) (2012). *Into the wild – Technology for open educational resources*. University of Bolton.
- [6] St. Lifer, E. (2018). Five Ed-Tech Elephants: From OER to equitable access, challenges to the digital revolution persist. *School Library Journal*, (1), 12.
- [7] Yergler, N. R. (2010). Search and Discovery : OER’s Open Loop. Retrieved from <http://openaccess.uoc.edu/webapps/o2/handle/10609/4852>
- [8] Simpson, M. D., & Twist, T. (2016). Curating Work-Integrated Learning: “Taking Care” of Disciplinary Heritage, Local Institutional Contexts and Wellbeing via the Open Educational Resources Movement. *Asia-Pacific Journal of Cooperative Education*, 17(1), 1–8.
- [9] Valiente, M.-C., Sicilia, M.-A., Garcia-Barriocanal, E., & Rajabi, E. (2015). Adopting the metadata approach to improve the search and analysis of educational resources for online learning. *Computers in Human Behavior*, 51, 1134–1141. <https://doi.org/10.1016/j.chb.2014.12.059>
- [10] Piedra, N., Chicaiza, J. A., López, J., & Tovar, E. (2014). An Architecture based on Linked Data technologies for the Integration and reuse of OER in MOOCs Context. *Open Praxis*, 6(2). <https://doi.org/10.5944/openpraxis.6.2.122>
- [11] Mouriño-García, M., Pérez-Rodríguez, R., Anido-Rifón, L., Fernández-Iglesias, M. J., & Darriba-Bilbao, V. M. (2018). Cross-repository aggregation of educational resources. *Computers & Education*, 117, 31–49. <https://doi.org/10.1016/j.compedu.2017.09.014>
- [12] Currier, S., Barton, J., O’Beirne, R., & Ryan, B. (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. *ALT-J, Research in Learning Technology*, 12(1).
- [13] Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2014). Metadata quality in learning object repositories: a case study. *The Electronic Library*, 32(1), 62–82.
- [14] Richter, T., & Veith, P. (2014). Fostering the Exploitation of Open Educational Resources. *Open Praxis*, 6(3). <https://doi.org/10.5944/openpraxis.6.3.139>
- [15] Belikov, O. M., & Bodily, R. (2016). Incentives and Barriers to OER Adoption: A Qualitative Analysis of Faculty Perceptions. *Open Praxis*, 8(3), 235–246.
- [16] Kinshuk, D., & Jesse, R. (2013). Mobile authoring of open educational resources as reusable learning objects. *The International Review of Research in Open and Distributed Learning*, 14(2), 28. <https://doi.org/10.19173/irrodl.v14i2.1535>
- [17] Carrion, C., Patricia, S., Morales, R., del Rocio, G., Pelaez, R., & Elizabeth, A. (2010). OER’s production cycle with social authorship and semantic tools (pp. 121–128). IEEE. <https://doi.org/10.1109/EDUCON.2010.5492588>
- [18] Allen, I. E., & Seaman, J. (2014). *Opening the Curriculum: Open Educational Resources in US Higher Education*, 2014. Babson Survey Research Group.
- [19] Ladbrook, J., & Prober, E. (2011). Information Skills and Critical Literacy: Where Are Our Digikids at with Online Searching and Are Their Teachers Helping? *Australasian Journal of Educational Technology*, 27(1), 105–121.
- [20] Rolfe, V. (2016). Web Strategies for the Curation and Discovery of Open Educational Resources. *Open Praxis*, 8(4), 297–312. <https://doi.org/10.5944/openpraxis.8.4.305>
- [21] Chicaiza, J., Piedra, N., Lopez-Vargas, J., & Tovar-Caro, E. (2014). Domain Categorization of Open Educational Resources Based on Linked Data. In *Knowledge Engineering and the Semantic Web* (pp. 15–28). Springer, Cham. https://doi.org/10.1007/978-3-319-11716-4_2
- [22] Dublincore.org (2012). Dublin Core Metadata Element Set, Version 1.1: Reference Description. [ONLINE] Available at: <http://dublincore.org/documents/dces/>. [Accessed 13 March 2018]

AUTHORS

First Author W. Jobe, is with University West in Trollhättan, Sweden. He is now with the Department of Informatics, (e-mail: william.jobe@hv.se).

Second Author S. Nilsson is with University West in Trollhättan, Sweden. He is with the department of Informatics, School of Media and Design (e-mail: stefan.nilsson@hv.se).

Third Author P. Bernhardsson is with University West in Trollhättan, Sweden. He is now with the Department of Informatics, where he works as a lecturer in digital media, focusing on web development (e-mail: patrik.bernhardsson@hv.se).

Fourth Author A. Bindslev Rask is with University College Nordjylland, Aalborg, Denmark (e-mail: andr@ucn.dk)

Manuscript received 25 May 2018.

Published as submitted by the author(s).