Getting Value from Open Educational Repositories

David Topps[1], Corey Wirun[2], Rachel Ellaway[2]

Corresponding author: Dr David Topps topps@ucalgary.ca
Institution: 1. University of Calgary, 2. OHMES, University of Calgary
Categories: Scholarship/Publishing, Teachers/Trainers (including Faculty Development), Technology

Received: 13/05/2019
Published: 31/05/2019

Abstract

Open Educational Repositories (OERs) have not seen the wide adoption that was predicted when they were introduced. A number of factors contribute to this but in this article, we explore some new tools and mechanisms that will create additional value from their use, for contributors, learners and organizations.

Keywords: open educational repositories; axiology; activity streams; xAPI

Background

Why have Open Educational Repositories (OERs), or learning object repositories, largely failed? Through the past decade, we have seen many of these set up and widely hailed for their virtues. ( But the shelves have remained largely empty (Kortemayer, 2013; Clark, 2015; Clements, 2016) and the corridors silent of footsteps. Looking at the literature, and discussing this with colleagues, has highlighted a number of salient factors:

What’s in it for me?

The lack of suitable reward structures is surely a factor. Altruism only goes so far. Yes, the appeal of OERs as contributing to the common good is very noble but suffers from the Tragedy of the Commons. In the daily lives and pressures facing academics, contributing to an OER has not put bread on the table or lines into one’s CV. Many of the factors proposed by Clark (Clark, 2015) really just relate to a wholesale skepticism, or cynicism, about the virtues of altruism. There must be a more tangible benefit to contributing.

Ease of contribution

For many OERs, the process of uploading an item is tedious, with far too many metadata items required as mandatory before the item is accepted. The frustrating interfaces for doing so were usually designed by consumers of the metadata (bibliometricians, archivists) and lacked easy mechanisms for uploading collections of similar objects. Having to manually enter arcane metadata codes and descriptors was seen as a poor return on investment (RoI) of time by most scholars.
Reusability
For the few objects that were deposited into OERs, they largely sat there ignored and unloved. This glaring lack of adoption was hidden by the lack of accessible reporting on usage and uptake from most OER platforms. But consider, as a simple common example of reuse, when you last asked a presenter for a copy of their slides. For some, this is simply another way of capturing the content and ideas, rather than relying on faulty note taking. But, for many, there is a slide or two in the deck which neatly portrays something that they also need to convey to an audience. For this, they really want just those slides, not the whole slide deck. Who amongst us has ever found, even with diligent searching, a complete presentation that meets the needs of our learners in its unaltered entirety? Reusable objects must be granular or modular.

Ideally, the reusable module should be plug-and-play. If the widget, or its content, is too entangled with its host habitat; or cannot be dropped into your own presentation format easily, then it is not likely to be reused. Simply using a hyperlink to the material is a common solution but also has the risk that the learner will hare off down some distraction pathway. (Figure 1)

Figure 1: the problem with Wikipedia

Discoverability
Simply finding useful material can be a challenge. Now, this is where librarians and archivists will smugly point out that this is why all that metadata is needed. And, in the world of traditional search strategies and Boolean syntax, this is true. But Google demonstrated, both in general search and with Scholar, that there are better ways to make things discoverable. (By making their metadata, but not the content, explorable by the web’s search spiders and robots, many OERs missed out on the value provided by the powerful capabilities of today’s search engines.)
For a short time, in the enthusiasm around the ‘semantic web’, we became excited by the possibility of semantic indexing, Resource Descriptor Format (RDF) coding, and the ability for these engines to “understand what I mean, not what I say” when asked questions. Semantic search has not gone away and remains a powerful technique but proponents in this field have realized that relying on authors and creators to help with supplying semantically structured metadata is fruitless. For the average author, there is even less perceived RoI to this approach.

Service Discovery

In the computer server software realm, Service-Oriented Architectures (SOA) promote the concept of Service Discovery to expose the devices or services that are offered by the system - absolving system stakeholders from the need for any advance knowledge of the capabilities offered by the system.

Adapted to the functionality of OER’s, a similar concept of Service Discovery could expose the data and capabilities of the OER to data consumers and software system integrators. As a start, a Service Discovery mechanism might publish a navigable and queryable set of ‘known’ metadata tags that the consumer can choose from. This provides a platform for potential metatag hierarchies to be exposed, which greatly facilitates complex ad hoc querying.

In a more ‘traditional’ SOA sense, an OER Service Discovery facility might also expose an Application Programming Interface (API) specification, using standard mechanisms such as the Web Services Description Language (WSDL). (Christensen et al., 2001) WSDL defines a mechanism that describes the methods and data structures that are available to be able to communicate with the system in an automated fashion. Software designers could then take the WSDL exposed by the OER and be able write software programs that automatically performs tasks such as data exchange or complex data analytics.

Quality Assurance

Much has been made of the need for quality. This also applied in the early days of the web: how was one to know that the web site or portal provided good quality and reliable material? Efforts such as the HON code (https://www.hon.ch/HONcode) remain well-meaning but unable to cope with the exponentially rising data floodwaters, and also susceptible to false claims and lack of enforcement. So, if we cannot measure or certify quality, can we measure engagement or impact?

For journals and journal articles, this is well established with impact factors, h-index and other such bibliometrics, a strong influence in judging the progress of an academic. We mentioned reward structures above – this has been a strong driver in the publish-or-perish academic game for the past two decades. But once we move beyond articles to other forms of scholarly output (Ellaway and Topps, 2017), things are much less clear. The efforts of Altmetrics (https://www.altmetric.com/) show great potential in this regard but these tend to focus on social media channels and newsworthiness, rather than educational impact.

There are even some oddities when comparing the two approaches. For example, there are many parts of the arguments raised by Clements (Clements, 2016) that we disagree with. From a social media perspective, this would not be a “Like”, but with citation indexing, there are now two citations of this article that were not there before. And on the long tailed curve seen for most articles, 2 citations is way more than most. (P. Bryan Heidorn, 2008) No such thing as bad publicity, as variously attributed to PT Barnum, Oscar Wilde or Mae West. (Martin, 2015)

Google and YouTube, world leaders in measuring what is popular and has impact, pay much attention to this and provide accurate reporting through Google Analytics. Perhaps we can learn from their attention to measuring activity streams and watching what people do, rather than how they respond to customer surveys.
Axiology vs Ontology

Much emphasis has been placed on ontology in OERs – the study of the form and nature of reality, and what can be known or described about it (Sandouk, 2015) – the metadata in this context. But continuing from the argument initiated above, should we not be more concerned with what people do with the contents of an OER?

Axiology is the study of value and what is worthwhile (Sandouk, 2015), and tends to be the forgotten cousin compared to ontology, epistemology and methodology as the pillars of scientific study. In our context here, axiology should be the driving force in how we evaluate the academic contributions and impact of learning objects, be they journal articles, datasets or slide decks.

In this regard, we should be more focused on paradata (data about usage, workflows etc) (Midgley and Gundy, 2011) than metadata. As noted above, measuring activity streams is the driving force behind the commercial fortunes of Google, Amazon and Facebook. The value proposition gained from such data is phenomenal: compare the yearly value, for one person, of this data to Google ($1200) with the cost of generating it ($0.005). (Madrigal, 2012)

If we want to close the loop in making OERs valuable as a resource, we need to create value to the authors in posting their materials there, and also to the institutions who pay their salaries but want to know what difference they are making. We must set up the infrastructure that measures such activity streams and shows us which learning objects are useful and have some impact. And we need to do this with an appreciation for the richness and interplay of such activities, and not just with a simplistic single number like an h-index. (Adams et al., 2019)

Closing such loops is a key focus of the PiHPES Project. (Topps, Ellaway and Greer, 2019) So, from that perspective, how can we integrate an OER into an infrastructure environment that can report on the axiology of its services, objects and activities? And how can we connect our various educational platforms to support such integration?

Functional requirements of an OER

Looking at the findings in various reviews, and looking at the principles determined in our own use cases arising in various educational scholarly projects over the past decade (Bamidis, Kaldoudi and Pattichis, 2009; Bratsas et al., 2009; Topps, Helmer and Ellaway, 2013; Topps et al., 2014; Dafli, Antoniou, Ioannidis, Dombros, Topps and P. D. Bamidis, 2015; Poulton, Woodham and Khavia, 2018), we suggest that the following requirements are the most essential:

- Easy data capture/upload – if the process of loading an object fails the RoI heuristics of the average scholar, there will be poor engagement and uptake. Automation or tools to support collections of similar objects are essential.
- Simple PURL generation – persistent URIs (the doi is the most widely known option) help to prevent link rot and make objects more respectable when cited.
- Easy citation generation – once an object is published, the OER should provide an easy way to display its citation in a variety of formats.
- Rich searchable metadata – more is better, except for the overworked, uploadig author. A wide range of possible fields, but with a minimum number of mandatory fields. Automatic capture from existing sources is highly desirable.
- Accepts wide rage of datatypes – the range of what is scholarly is broad. (But in Precision Education
and big data analytics, the OER should not constrain file sizes or storage needs unnecessarily.

- Flexible affiliation metadata – not all sub-units of an institution, or all collections of scholars, can be grouped according to standard structures like departments or faculties. OHMES, as a horizontal structure within CSM, and with its mandate to collaborate broadly outside of its host university, is particularly sensitive to this limitation. This affects activity reporting and Altmetrics.

- Associated files, guides – may learning objects and datasets require associated documents. These should be easily attached to the primary item, along with metadata as needed.

- Paradata – usage or activity metrics to inform the axiology of the contents. We recognize that the paradata generated when the objects are embedded into new materials will be more important but OER paradata is still helpful. A common mechanism to collate such paradata would be ideal.

- Version control – simple version control may be helpful. However, for objects that change frequently (which can be a notable barrier for the doi system), where it is important that changes between versions are well documented, a system like GitHub is superior.

**Lessons learned from past projects**

Over the past decade, in various projects in which we have collaborated, we have noted a number of applications and standards that, while looking initially promising, have proven to be barriers or unhelpful.

Mechanisms for integrating information flow across multiple educational applications are not new. SCORM (Dodds, 2004) dates back to 1999 and was a game changer, compared to the anarchy that preceded it. Based on the needs of the time, Phil Dodds and his colleagues crafted a remarkably strong standard to enable learning management systems and associated applications to share common objects. But SCORM, while widely requested, has not been extensively implemented. It appears to be too rigid in its ability to connect systems. A change in one system usually breaks SCORM compatibility, which is referred to as tight-coupling. And changes are common in information systems, as we all know. Moreover, little actual data was captured about what was done in the modules. All that was passed back to the parent system was that the learner did or did not complete the module, along with a single pass mark. We are pleased to note that more modern alternatives, such as the Experience API, convey much richer data about who did what when. (Lindert and Su, 2016)

The mEducator (Dafli, Antoniou, Ioannidis, Dombros, Topps and P. Bamidis, 2015), and associated projects, looked initially promising in their use of semantic metadata to more fully describe learning objects in an OER or in other applications such as OpenLabyrinth. After looking at a variety of RDF-based vocabularies, the group settled on SKOS [https://www.w3.org/2004/02/skos/](https://www.w3.org/2004/02/skos/) and integrated it successfully into their infrastructure. Although the semantic search and the SKOS integration were technically powerful, the main barrier appeared to be in adoption. Object authors were expected to tag their materials with RDF coding, a somewhat tedious process. The initial emphasis was on trying to make the RDF coding robust and accurate, which makes the annotation process even more labour-intensive. Our own subsequent explorations (Topps, 2015) in this area suggested that a simple, automated annotation tool, which provided "dirty" inexact coding, was still sufficient to support useful semantic search functionality but this came too late in the project. The systems remain in place for others to explore, but have seen little uptake.

The Learning Registry [https://github.com/LearningRegistry/LearningRegistry/wiki](https://github.com/LearningRegistry/LearningRegistry/wiki) was a promising initiative, supported by the US Department of Education until Sept 2018. In its design, it was an improvement on previous referatories, such as MERLOT, and used open standards and APIs so that it would be much easier to seed the OER with materials, as we have suggested earlier in this paper. One drawback, as seen in other referatories, is that there
was no significant capacity to store large datasets or objects. This renders the system vulnerable to link rot and the persistent availability of the source objects. It is too early to foretell its future but the lack of widespread adoption is concerning.

The Kritikos project at the University of Liverpool (Bullough, Mannis and Green, 2014) has some interesting facets, which made it appear very promising when introduced in 2012. Although it centered around a visual search engine, we were more interested in its functions that were designed to support and facilitate user-generated content and contributions. Students were encouraged to send in links to data that they themselves had found to be useful material in any given course. Again, as a demonstration of the Tragedy of the Commons, the level of contribution by learners very rapidly dropped off after some initial enthusiasm. We suspect that the lack of reward, either intrinsic or extrinsic, for the contributors, was a significant factor. The system is still in place but shows little sign of use.

PiHPES proposed plumbing

Our Precision in Health Professions Education Scholarship (PiHPES) project (https://dataverse.scholarsportal.info/dataverse/pihpes) is exploring a number of tools and approaches that can support the integration of data streams across educational platforms.

In particular, we are using activity metrics across many of our tools and systems, collected into a common set of Learning Record Stores, as a means to capture what our teachers and learners do in a variety of contexts. We are using Experience API (xAPI) formatted activity statements that are generated during the use of a number of applications, modules and widgets within the PiHPES platforms. The functionality of these objects, and the activity metrics that are generated by these objects, are abstracted from the context of their use.

Consistent with the recommendations made earlier in this paper, we are making the modules and widgets available as reusable, portable, learning objects, with a finer granularity so that they can be easily redeployed as small components, rather than whole courses or scenarios. This will require a mechanism that can support the aggregation, annotation and deployment of these objects. Originally, we had simply planned to make these scoped objects available within their parent applications, to be exported as the need arose.

But given the considerations raised in this document, and by our previous experiences with such shared object methods, it now appears to be more advisable to build upon previous lessons learned, by us and others, and to take a more sustainable approach in making these reusable, scoped objects more accessible to both teachers, learners and staff.

Part of the need will be to find ways to reward contributors of these objects for their efforts. Altruism only goes so far. By instantiating these objects as items in an OER, that supports the functional requirements listed above, we will be able to establish mechanisms where authors and contributors will receive academic credit for their efforts. They will also be able to analyse the activities in which their modules, scenarios or widgets are employed, in order to close the feedback loop and improve the quality of the contributions.

Does a Dataverse do it?

Over the past few years, we had explored a number of different OERs and other mechanisms for making such objects more available. Indeed, at the annual Medbiqitous (www.medbiq.org) meeting two years ago, there was a general call to the groups working on virtual scenarios to establish a global common store of such scenarios to
facilitate sharing. There is a long history of such shared cases (Smothers, Ellaway and Balasubramaniam, 2008; Bamidis, Kaldoudi and Pattichis, 2009; Poulton, Woodham and Khavia, 2018) but the underlying repositories lacked many of the suggested functions.

Our group recently became aware of the Harvard Dataverse (https://dataverse.harvard.edu/dataverse/harvard), an OER that is publicly available and that has many of the functions mentioned. Some early work with this OER has been very promising but there have also been some significant limitations:

- Heavy usage has made the system very slow in response and unreliable.
- Heavy use has also created some challenges for their tech support teams, who were initially quite responsive.
- Data is stored on USA servers, which gives discomfort to some and raises jurisdictional issues for some datasets.

Use is free of charge but then we are liable to the Pigs-in-a-barn effect (Widder, 2010) – see Figure 2.

Figure 2: the Free Model

We were delighted to discover that there is an all-Canadian version available through the University of Calgary libraries (https://dataverse.scholarsportal.info/dataverse/ohmes). The Scholars Portal Dataverse (https://dataverse.scholarsportal.info) is hosted by the University of Toronto, run on Canadian cloud services, and managed by a consortium of Canadian institutions. It uses the same software and is not so heavily loaded, and therefore much more responsive. It is fully supported by UCIT.

In general, in our explorations of the Dataverse platform so far, we have noted the following:

- We can use its functions to support claims for academic credit
It provides all the metadata functionality that we specified above
It provides doi ad citation services
It accepts a full range of datasets and sizes
It ca be federated, making collaboration with non-Canadian project teams feasible
It provides support for automated uploads of collectios of similar objects

While it does provide some simple activity metrics, we feel that this can be improved. This can be further explored by the OHMES technical experts.

Technical Report
Some of the topics and technical details touched upon in this paper require much greater explication. We have created an associated Technical Report to supplement this article. (Topps, Wirun and Ellaway, 2019) The Technical Report is very specific to the PiHPES Project. We hope that the observations, lessons, and recommendations contained in this paper are more generalizable to other institutions that are also facing data integration challenges.

Take Home Messages
OERs provided little value to their contributors previously. Using activity metrics for both producers and consumers should mitigate this deficiency.

Assistance with batch uploads of similar datasets should enable greater usage and reduced effort.

Notes On Contributors
Prof David Topps is Medical Director of OHMES (Office of Health & Medical Education Scholarship) for the Cumming School of Medicine at the University of Calgary. https://orcid.org/0000-0003-3593-544X
Corey Wirun is Senior Analyst at OHMES (Office of Health & Medical Education Scholarship) for the Cumming School of Medicine at the University of Calgary. https://orcid.org/0000-0002-4355-9790
Prof Rachel Ellaway is Director of OHMES (Office of Health & Medical Education Scholarship) for the Cumming School of Medicine at the University of Calgary. https://orcid.org/0000-0002-3759-6624

Acknowledgements
- Gretchen Greer for project management, data collection and analysis, operational support
- Maureen Topps, former Senior Associate Dean Education, for funding support and consulting on organizational readiness
- Bev Adams, Senior Associate Dean Education, for funding and operational support

Bibliography/References
Adams, J., Mcveigh, M., Pendlebury, D. and Szomszor, M. (2019) Profiles, not metrics. Available at:
Bamidis, P. D., Kaldoudi, E. and Pattichis, C. (2009) ‘mEducator: A Best Practice Network for Repurposing and Sharing Medical Educational Multi-type Content’, in *IFIP Advances in Information and Communication Technology*. Springer, Berlin, Heidelberg, pp. 769–776. https://doi.org/10.1007/978-3-642-04568-4_78

Bratsas, C., Kapsas, G., Konstantinidis, S., Koutsouridis, G., et al. (2009) ‘A semantic wiki within moodle for Greek medical education’, in *2009 22nd IEEE International Symposium on Computer-Based Medical Systems*. IEEE, pp. 1–6. https://doi.org/10.1109/CBMS.2009.5255417

Bullough, T., Mannis, A. and Green, A. (2014) *Kritikos*.

Christensen, E., Curbera, F., Meredith, G. and Weerawarana, S. (2001) *Web Service Definition Language (WSDL)*. Available at: https://www.w3.org/TR/2001/NOTE-wsdl-20010315 (Accessed: 13 May 2019).

Clark, D. (2015) *Donald Clark Plan B: 10 reasons why OER fails. CAVE dwellers (Colleagues Against Virtually Everything)*, *Donald Clark Plan B*. Available at: http://donaldclarkplanb.blogspot.com/2015/05/is-closed-mindset-of-open-educational.html (Accessed: 8 May 2019).

Clements, K. (2016) *Why Open Educational Resources Repositories Fail The Contribution of Quality Approaches to the Success of Repositories*. University of Jyväskylä, Finland. Available at: https://jyx.jyu.fi/bitstream/handle/123456789/49262/978-951-39-6588-4_vaitos_20160416.pdf (Accessed: 8 May 2019).

Dafli, E., Antoniou, P., Ioannidis, L., Dombros, N., Topps, D. and Bamidis, P. D. (2015) ‘Virtual patients on the semantic web: A proof-of-application study’, *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.3933

Dodds, P. (2004) *SCORM Overview, Advanced Distributed Learning*. Available at: https://www.adlnet.gov/adl-research/scorm/

Ellaway, R. and Topps, D. (2017) ‘METRICS: a pattern language of scholarship in medical education’, *MedEdPublish*, 6(4). https://doi.org/10.15694/mep.2017.000199

Haag, J. (2015) *xAPI Overview - ADL Net, ADL Net*. Available at: https://www.adlnet.gov/xAPI (Accessed: 29 May 2017)

Heidorn, P. B. (2008) ‘Shedding Light on the Dark Data in the Long Tail of Science’, *Library Trends*. Johns Hopkins University Press, 57(2), pp. 280–299. https://doi.org/10.1353/lib.0.0036

Kortemayer, G. (2013) ‘Ten Years Later: Why Open Educational Resources Have Not Noticeably Affected Higher Education, and Why We Should Care | EDUCAUSE’, *Educause Review*. Available at: https://er.educause.edu/articles/2013/2/ten-years-later-why-open-educational-resources-have-not-noticably-affected-higher-education-and-why-we-should-care (Accessed: 8 May 2019).

Lindert, L. and Su, B. (2016) ‘The Evolution of SCORM to Tin Can API: Implications for Instructional Design’,
Madrigal, A. C. (2012) *How Much Is Your Data Worth?*, *The Atlantic*. Available at: http://www.theatlantic.com/technology/archive/2012/03/how-much-is-your-data-worth-mmm-somewhere-between-half-a-cent-and-1-200/254730/ (Accessed: 23 August 2016).

Martin, G. (2015) ‘There is no such thing as bad publicity’ - the meaning and origin of this phrase, *The Phrase Finder*. Available at: https://www.phrases.org.uk/meanings/there-is-no-such-thing-as-bad-publicity.html (Accessed: 9 May 2019).

Midgley, S. and Gundy, V. (2011) *Paradata (learning resource analytics)*, *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Paradata_(learning_resource_analytics).

Poulton, T., Woodham, L. and Khavia, S. (2018) *Waves Network*. Available at: http://www.wavesnetwork.eu/index.php#page-home (Accessed: 8 May 2019).

Sandouk, M. (2015) *What is axiology and how does it relate to ontology and epistemology?*, *CES PhD Support Group*. Available at: https://blogs.warwick.ac.uk/cesphd/entry/what_is_axiology_1/ (Accessed: 8 May 2019).

Smothers, V., Ellaway, R. and Balasubramaniam, C. (2008) ‘eViP: sharing virtual patients across Europe.’, *AMIA Annu Symp Proc*. United States, p. 1140.

Topps, D. (2015) ‘Education research keynote at MEI 2015’, in Bamidis, P. (ed.) *Medical Education Innovations*. Thessaloniki, Greece.

Topps, D., Ellaway, R. and Greer, G. (2019) *PiHPES Project: broadened perspectives on learning analytics*. Calgary. https://doi.org/10.5683/SP2/VDDGG3

Topps, D., Helmer, J. and Ellaway, R. (2013) ‘YouTube as a platform for publishing clinical skills training videos’, *Academic Medicine*. https://doi.org/10.1097/ACM.0b013e31827c5352

Topps, D., Taenzer, P., Armson, H., Carr, E., et al. (2014) *DynIA: Dynamically Informed Allegories. (final report)*. Calgary, AB, Canada. Available at: http://dspace.ucalgary.ca/handle/1880/50360

Topps, D., Wirun, C. and Ellaway, R. (2019) *OER integration for PiHPES - PiHPES Dataverse*. Calgary. https://doi.org/10.5683/SP2/VIHYNC

Widder, O. (2010) *The 'Free' Model, Geek and Poke*. Available at: http://geek-and-poke.com/geekandpoke/2010/12/21/the-free-model.html (Accessed: 13 May 2019).

**Appendices**

None.
Declarations

The author has declared that there are no conflicts of interest.

This has been published under Creative Commons "CC BY 4.0" ([https://creativecommons.org/licenses/by-sa/4.0/](https://creativecommons.org/licenses/by-sa/4.0/))

Ethics Statement

No study subjects involved.

External Funding

This paper has not had any External Funding

AMEE MedEdPublish: rapid, post-publication, peer-reviewed papers on healthcare professions’ education. For more information please visit [www.medepublish.org](http://www.medepublish.org) or contact medepublish@dundee.ac.uk.