

Examining the Reuse of Open Textbooks



John Hilton III and **David A. Wiley**
Brigham Young University, USA

Neil Lutz
Rutgers University, USA

Abstract

An important element of open educational resources (OER) is the permission to use the materials in new ways, including revising and remixing them. Prior research has shown that the revision and remix rates for OER are relatively low. In this study we examined the extent to which the openly licensed *Flat World Knowledge* textbooks were being revised and remixed. We found that the levels of revision and remix were similar to those of other OER collections. We discuss the possible significance and implication of these findings.

Keywords: Open educational resources; remix; reuse; open textbooks; electronic textbooks; Flat World Knowledge; open access

Only a decade ago textbooks existed only as a physical resource. When two roommates shared a textbook, if one roommate was using it the other could not. However, some resources (like textbooks) are now available in digital formats. Two people (or a million, for that matter) could all be online and using the same digital textbook at the same time. Because the resource is digital, it is more accessible.

One way to further increase the accessibility of the digital textbook is to make it an *open educational resource* (OER). By OER it is meant that others can reuse and/or modify the digital textbook (a more in-depth discussion of what it means to be an OER will follow). Assuming the textbook was written in English, the only people who will be able to benefit from the textbook are those who read English. But suppose the creators of the digital textbook had made it an OER (i.e., they allow others to make changes to it) and someone translates the textbook and makes it available in Chinese. The OER digital textbook is now more accessible to others—making the textbook increasingly accessible. But what if nobody had made any modifications to the textbook? It could be argued that any time or expense spent in openly licensing the textbook—such that it could be modified—would have been wasted

since this capacity was never utilized.

Today there are thousands of OER available for others to reuse and modify. In the present study we examine the question as to how many people are taking advantage of the open licensing feature of OER textbooks that allows these resources to be revised and remixed.

Review of Literature

Open educational resources have become an important topic of research. According to the Organisation for Economic Co-operation and Development (OECD), “The definition of OER currently most often used is ‘digitised materials offered freely and openly for educators, students and self-learners to use and reuse for teaching, learning and research’” (OECD, 2007, p. 10). These resources, made freely available on the Web, provide those with Internet access the opportunity to tap into a variety of educational materials at no cost.

The Cape Town Open Education Declaration (2007) states,

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.

Similarly, India’s National Knowledge Commission has written that stimulating “the development and dissemination of quality Open Access (OA) materials and open educational resources (OER) through broadband Internet connectivity” would be one of the most effective ways to disseminate educational resources in India (National Knowledge Commission, 2007, p. 51).

The *open* in OER is typically understood to mean that the resource is freely available for others to use in different contexts (McMartin, 2008). But the word *reuse* can describe several different types of use, and not everyone agrees on what reuse means (Downes 2007, Hylén, 2006). Hilton, Wiley, Stein, and Johnson (2010) describe four Rs that clarify nuances in what reuse means.

- *Reuse*—The most basic level of openness. People are allowed to use all or part of the work for their own purposes (e.g., download an educational video to watch at a later time).
- *Redistribute*—People can share the work with others (e.g., email a digital article to a colleague).
- *Revise*—People can adapt, modify, translate, or change the work (e.g., take a book written in English and turn it into a Spanish audio book).

- *Remix*—People can take two or more existing resources and combine them to create a new resource (e.g., take audio lectures from one course and combine them with slides from another course to create a new derivative work).

An important element of OER is permission to use the materials in new ways. In fact, D’Antoni (2009) cited the William and Flora Hewlett Foundation, a key donor for OER, as stating that OER are “resources that reside in the public domain or have been released under an intellectual property license that permits their free use or re-purposing by others” (p. 4).

How an individual licenses an OER significantly affects its openness. In many jurisdictions, all creative works are automatically copyrighted; therefore, they are legally “closed” unless the author takes explicit steps to open them (Lessig, 2004).

One way to take the steps necessary to open creative works is to use a Creative Commons license. Creative Commons provides several licenses to help creators of content license their work in ways consistent with their desires for openness (e.g., Creative Commons, 2009; Bissell, 2009). These licenses generally allow people to take digital objects (like the textbook) and reuse them in other settings. For example, elementary school teachers could put the textbook on their personal Web sites and point their students to those sites.

There are certainly benefits to this level of openness, which allows for reuse. However, Windle, Wharrad, McCormick, Laverty, and Taylor (2010) pose questions about reuse:

This current OER movement is opening up repositories to be populated with all sorts of materials but does it matter if OER is concerned with simply filling repositories with resources, and have institutions fulfilled their ‘public’ obligation by doing so? Are they also responsible for whether people actually use it and for supporting and monitoring whether they do so? (n.p.)

They argued that reuse is an important part of OER, but their research stopped short of quantifying how much OER are being reused. Similarly, Koper (2003) asks, “Do users actually make use of our systems? In reality, do they share and reuse objects?” (47). Thus a key question is, given the fact that OER are being created and licensed for revision and remix, is anybody revising and remixing?

Although little research has been done in this area, the literature indicates that when OER are used, they are typically adopted wholesale (reused), and *not* revised or remixed (Duncan, 2009; Petrides, Nguyen, Jimes, & Karaglani, 2008). This distinction is important. As with the textbook example given in the introduction, reuse is simply taking the OER and using it as is. Revision would be modifying the textbook in some way (e.g., translating it) in order to increase its usefulness. A textbook could also be revised by having professors remove chapters that were not germane to their classes or add in material they wanted to

cover. Such a textbook could be remixed by taking preexisting openly licensed images, videos, or writings and merging them into an existing textbook.

Allowing for revisions and remixing can potentially enhance the reach of a work (Bissell, 2009). And this does happen in some cases. For example, Hilton (2009) found that in one instance a book was translated into several languages and file formats as a result of being openly licensed. However, this is atypical. Duncan (2009) found that in one OER collection the rates of revision or remix were quite low. Of 3,519 modules created and used (an important distinction, since 32.6% of created modules were never used), only 2.98% of modules were translated, and only 2.87% were modified in some other way. Thus, according to the definitions of revision and remix we have set forth, approximately 6% of OER in this collection were ever revised or remixed.

Petrides et al. (2008) found that while a large number of what they termed “reuses” occurred on the same OER collection studied by Duncan, they also stated that “88 percent [of these reuses] . . . involved author users manipulating their own content” (p. 350). People revising their own work does not seem to fit the spirit of revision or remix embodied in the idea of open educational resources. These findings challenge the idea that people are eager to take OER that others have created and revise or remix them for their own purposes. Vuorikari and Koper (2009) found a reuse rate of approximately 20%. However, their definition of reuse was significantly more inclusive than those used by Petrides et al. (2008) and Duncan (2009). Van der Baaren and de Vried (2010) tell of pilot efforts to increase the reuse of OER; however, they also state that it is too early to report on the outcomes of these pilot studies.

Of course, none of these studies can fully take into account remixing that was done outside of the OER collection. For example, had a teacher taken content from the collection offline, remixed it with other content, and posted it on his or her own Web site, it would not have been counted as a remix in these studies. This unobservable exercise of the 4R permissions has been called “dark reuse” (Wiley, 2009).

The Present Study

We examined the extent to which professors reused, revised, and remixed OER in the form of textbooks published by *Flat World Knowledge*. Founded in 2007, Flat World Knowledge (FWK) makes all of its textbooks available with an open license (Creative Commons BY-NC-SA). Digital versions of these books are freely available to anyone with an Internet connection. Flat World Knowledge makes its money from sales of print versions of textbooks (some students still prefer the traditional book) as well as other supplemental materials (like digital flash cards, study guides, etc.).

A key measure used by FWK is the “adoption” of textbooks. *Adoption* means that a professor formally decides to use a FWK textbook in his or her classrooms for a given semester. For example, if an English professor taught three sections in the fall semester and four sections in the spring semester using a FWK book, this teacher would have accounted for

seven adoptions.

FWK has a proprietary system that allows professors to revise and remix the content of the book they will use. Professors can make changes such as deletions, reorderings of material, additions of a professor's own material, and additions of FWK or other openly licensed materials. According to the definitions of 4 Rs set forth by Hilton et al. (2010), deletions, reordering, and other adaptations would be classified as revisions, and additions of other pre-existing openly licensed materials (from FWK, YouTube, etc.) would be classified as remix.

Our specific purpose was to determine the following: (a) how many of the total adoptions did any kind of revising or remixing; (b) in cases where revisions or remixes occurred, what kinds of changes were made; and (c) if there were any consistent patterns or trends in the remixes.

Measuring Revisions

In order to measure the extent of a revision, we used a multistep approach. First, we matched all corresponding sections of the custom book and the original book. Flat World Knowledge's customization framework allows for rearrangement at the table of contents level: chapters may be reordered within the book, sections may be reordered within a chapter, and sections may be moved between chapters. There is no tool provided for reordering the content within sections, so sections were treated as 'atomic' from the rearrangement perspective. As sections are moved and edited, they maintain a unique ID that can be used to find matches.

If a section in the original had no match in the custom, then the section had been deleted, and its entire length was counted toward the *delete score*. If a section in the custom had no match in the original, then the section had been added, either directly from the customizer or as a remix from other books in the Flat World Knowledge catalog. The length of the section then counted toward the *add score* or *external add score*, respectively. If an existing paragraph was cut and replaced by a new paragraph this counted as an addition and a deletion.

When a match was found, we directly compared the pair to find shared text. Given two sections, we measured the similarity between them by finding the length of the *longest common subsequence*, or LCS, which represents the text that is contained, in order, in both sections. Here is an example:

```
First, this is the original sentence.
And this is a revised sentence.
```

These have lengths of 37 and 31, ignoring spaces. We match them up as well as possible by adding spaces.

```
First, this is the original sentence.
And   this is a   revised sentence.
```

The LCS here is “this is r i sentence.” because these characters appear in that order in both sentences. This has length 16. To get from the first sentence to the second, one could delete everything that is in the first but not the second (“First, the o ig nal”, length $37 - 16 = 21$) to get down to the LCS, then add everything in the second but not the first (“And a ev sed”, length $31 - 16 = 15$). So the add score for this minirevision is $15 / 31 = 48\%$, and the delete score is $21 / 31 = 68\%$.

Toolkit

FWK data regarding the contents of textbooks, customizations, and adoptions is stored in MySQL databases (see <http://dev.mysql.com/>) and privately accessible through their proprietary BookServer API. We extracted and processed data using Python (see <http://python.org/>), relying heavily on difflib (see <http://docs.python.org/library/difflib.html>) and Microsoft Excel for our analysis. Graphs were created using the R statistical environment (see <http://www.r-project.org/>) and the ggplot2 graphing library (see <http://had.co.nz/ggplot2/>).

Results

Between fall 2009 and summer 2011, there were 3,304 adoptions of FWK textbooks, 247 (7.5%) of which were adoptions of custom books that differed from the original.¹ In addition, we examined this subset of books (referred to as *custom books*) to see how these custom books were altered.

Analysis of Customization Behavior

Four types of customizations were available to users of the FWK customization system: additions, deletions, reorderings, and remixes.

Additions.

An addition occurred when a user added material to a book by inserting one or more new chapters, sections, paragraphs, or characters in an existing paragraph. The total number of custom books with additions was 41 (out of 247), or 16.6% of all custom books. Five custom books included only additions and no other customizations (2% of custom books). The amount of material added to a book ranged from 0.1% to 6.8% of the total content of the

¹ Interestingly, some of the books customers took through the customization process emerged from that process unscathed. Twenty-eight of the adopted books marked “custom” in the database were identical to the original books. One example of how this might happen would be a faculty member deciding to customize a book, opening it in the customization engine, and then deciding on further inspection there’s nothing he actually wants to change and hitting ‘save.’ These books are excluded from our analysis of customization behavior. There were also several customized books that were never adopted. One example of how this might happen is a faculty member playing around with the customization engine just to see how it works. These books are also excluded from our analysis of customization behavior. This in itself is an important finding of this study.

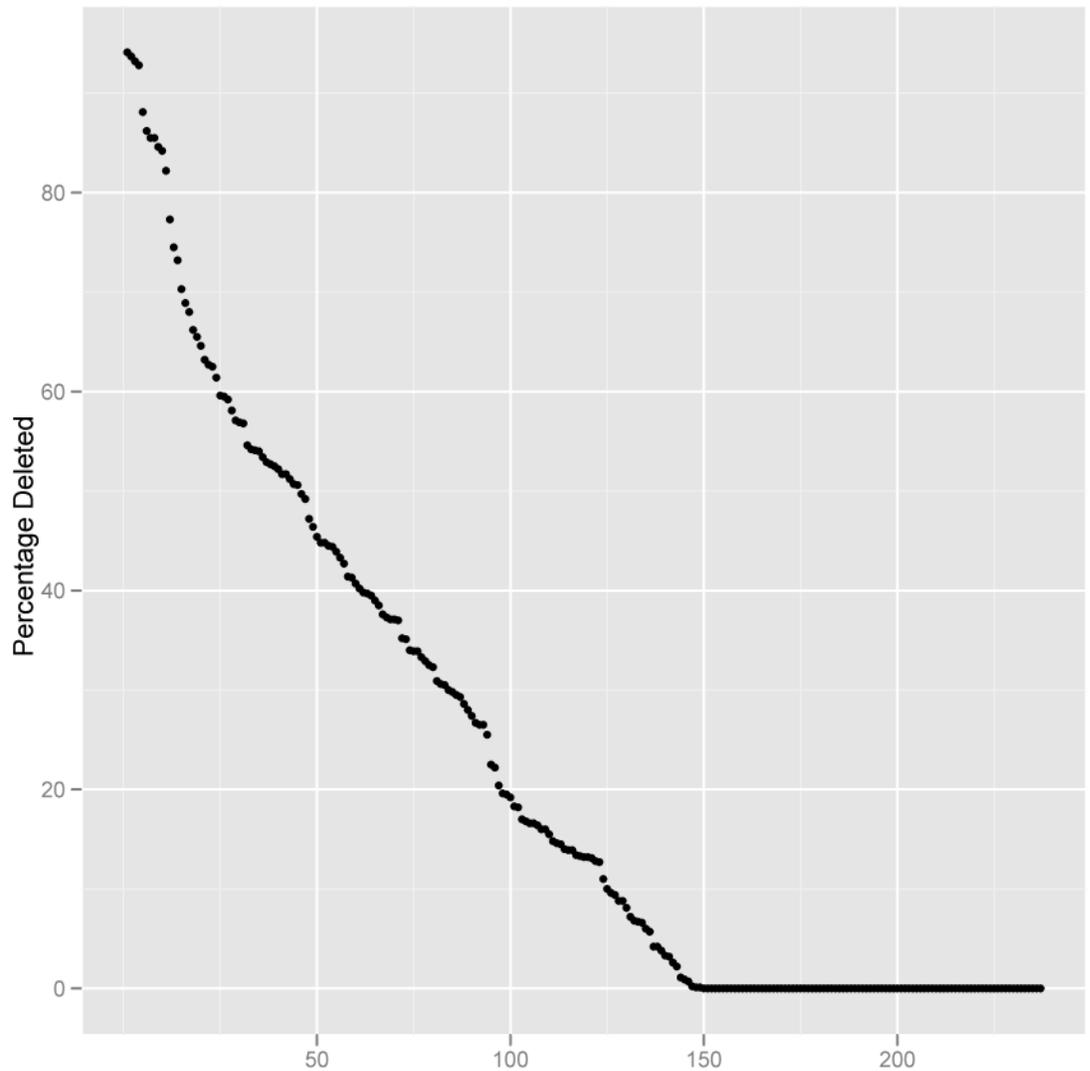


Figure 2. Amount of material deleted to custom books.

Reorders.

A reorder occurred when a user changed the sequence in which chapters appeared in the book or the sequence in which sections appeared in a chapter. The total number of custom books with reorders was 100 (out of 247), or 40.5% of all custom books. Ten custom books included only reorders and no other customizations (4.1% of custom books).

Remixes.

A remix occurred when a user imported content from another FWK book into the custom book. The total number of custom books with remixes was four (out of 247), or 1.6% of all custom books. No custom books included only remixes and no other customizations. The amount of material from other FWK books added to a book ranged from 12.3% to 79.6% of

the total content of the custom book. It is important to note that the remix rate could have been underestimated because it was assumed that all additions were of the professors' own creation. However, a professor could have copied and pasted material from a non-FWK textbook, and our methodology would not have detected this as a remix.

Patterns in Customization Behavior

We found a strong relationship between user customization behavior and the simplicity of accomplishing the different types of customizations. The simplest customization task, deletion, requires only a single click. Reordering the elements in the table of contents requires dragging and dropping. Adding original material to a book requires clicking and typing. Remixing material from another FWK book cannot be done directly by users and requires support from the publisher. The simpler a customization is to perform, the more customers perform it.

Table 1

Customization Types

Customization type	Number of custom books	Percentage of custom books
Deletions	149	60.32%
Reorderings	100	40.49%
Additions	41	16.60%
Remixes	4	1.62%

The same pattern holds for customized books including only one type of customization.

Table 2

One Customization Type Only

Customization type	Number of custom books	Percentage of custom books
Only Deletions	56	22.67%
Only Reorderings	10	4.05%
Only Additions	5	2.02%
Only Remixes	0	0.00%

Deletions were both the simplest customization to make and the most common customization. Some customizers deleted aggressively.

Table 3

Deletions

Amount of original book deleted	Number of custom books	Percentage of custom books
At least 25% deleted	94	38.06%
At least 50% deleted	45	18.22%
At least 75% deleted	12	4.86%

A strong relationship exists between the length of a book and the average number of chapters deleted from the book. A Pearson’s product-moment correlation was calculated to reveal the relationship between the number of chapters in a book and the average number of entire chapters deleted from custom versions of the book, resulting in a correlation of 0.71. Figure 3 demonstrates this relationship.

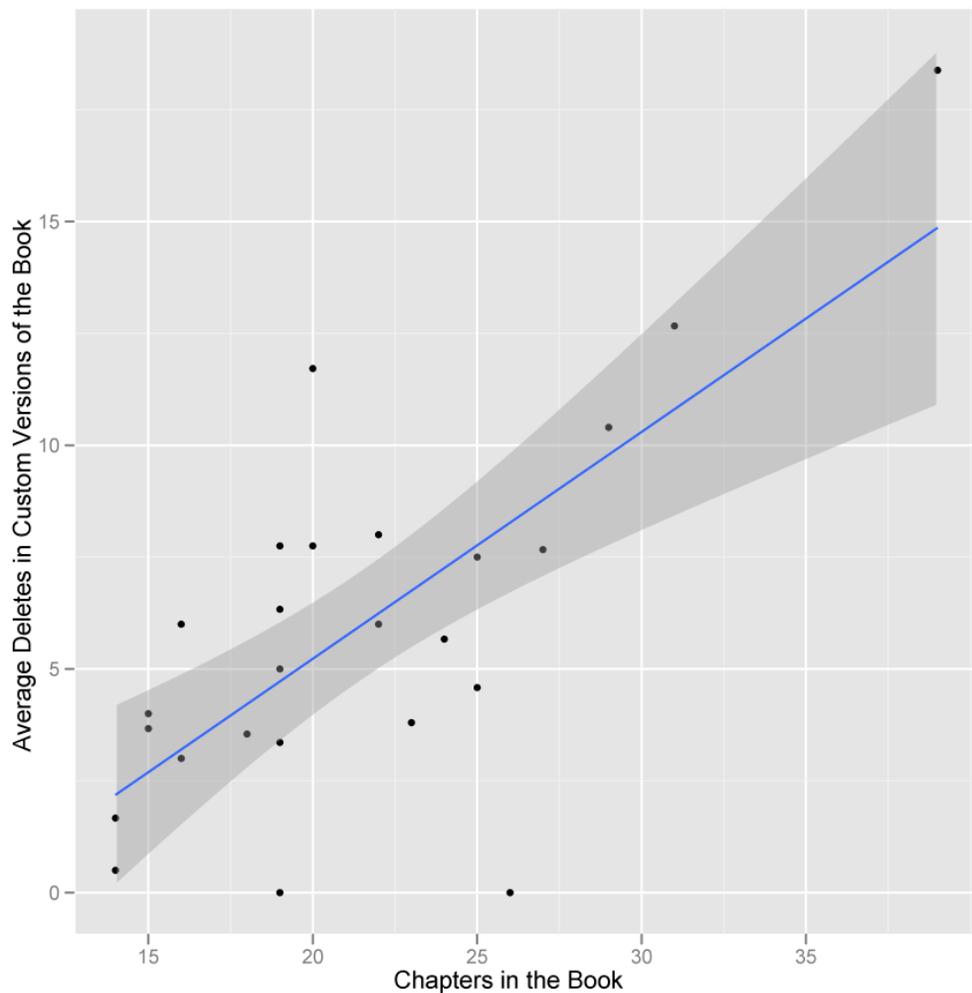


Figure 3. Relationship between book length and material deleted from custom books.

This propensity to remove entire chapters is particularly interesting given that the removal

does not lower the price of the online book (which is free) or the print book (which has a fixed cost, regardless of how it is customized). Given the lack of financial incentive to engage in the book customizing process, one interpretation of this pattern of behavior is that faculty who customize books do so because they prefer shorter textbooks tailored specifically to their course designs.

Discussion

As with Duncan (2009), we found that the rates of revision and remix were relatively low. Only 7.5% of textbook adoptions over a two-year period were adoptions of custom books. This indicates that while the ability to revise and remix sounds exciting, the number of those who take advantage of this opportunity is relatively small.

One tentative key finding of the study is that the easier it was to make a change, the more changes of that type were made. Of the different kinds of changes that can be made, the easiest to make is a deletion. And as was seen, deletions were by far the most common kind of change made to the FWK text. While we cannot state that deletions were made because they were easiest to make (perhaps they are the type of change that people wanted to make the most), this could have been a part of the reason. If this is the case, it stands to reason that the easier it becomes to revise, the more revisions will take place. While beyond the scope of the present study, it would be interesting to do additional studies to determine whether OER in easy-to-edit formats (e.g., text) are revised more frequently than OER in formats that are more difficult to edit (e.g., video).

Another finding was that not only were deletions the most common type of revision made, but they were also sometimes used heavily. Specifically, 38% of customizations deleted at least 25% of the original textbook and an additional 18% deleted at least 50%. Why did such steep deletions take place? While such deletions would have made the textbook shorter, the cost of the textbook to the students would have remained the same. In other words, there was no financial incentive for deletions to be made. We speculate that professors made deletions in order to remove material that was irrelevant to their class. It is also possible that some professors were using the FWK textbook as a supplement to another textbook, and thus only wanted to use a smaller part of it.

Limitations and Future Research

The primary limitation of the current study is that we only had access to modifications made within FWK's editing program and could not examine "dark reuse" of FWK content. It is possible that people revised or remixed FWK books outside of this system. For example, one of the authors of this article is currently revising the FWK project management textbook to focus on project management in the instructional design field. Consequently, this study undercounts (potentially significantly) the amount of actual remix that has taken place. In addition, FWK has recently added a new feature to its system in which people who customize textbooks can make their revised version of a textbook available for sale and

receive a portion of the profits from the sales of these remixed books. This incentive may encourage more professors to revise and remix textbook content.

Another limitation of this study is that it included only one collection of OER (those published by FWK). While our results are not significantly different from those found by Duncan (2009) and Petrides et al. (2008), additional research should examine other collections of OER to determine if these emerging trends in reuse, revision, and remixing hold in other venues.

Further work should be done not only to determine the extent to which OER are being remixed, but also how important this remixing is to OER in general. It may be that other outcomes of OER (simple reuse, decreased costs, increased access) overshadow the overall importance of how much remixing takes place.

References

- Bissell, A. (2009). Permission granted: Open licensing for educational resources. *Open Learning: The Journal of Open and Distance Learning*, 24(1), 97–106.
- Cape Town Open Education Declaration. (2007). *Cape Town open education declaration: Unlocking the promise of open educational resources*. Retrieved from <http://www.capetowndeclaration.org/read-the-declaration>
- Creative Commons. (2009). *Creative Commons licenses*. Retrieved from <http://creativecommons.org/about/licenses/meet-the-licenses>
- D’Antoni, S. (2009). Open educational resources: Reviewing initiatives and issues. *Open Learning, The Journal of Open and Distance Learning*, 24(1), 3–10.
- Downes, S. (2007). Models for sustainable open educational resources. *Interdisciplinary Journal of Knowledge and Learning Objects*, 3, 29–44.
- Duncan, S. M. (2009). Patterns of learning object reuse in the Connexions repository. *All Graduate Theses and Dissertations. Paper 423*. Retrieved from <http://digitalcommons.usu.edu/etd/423>
- Hilton, J. (2009). From PDF to MP3: Motivations for creating derivative works. *First Monday*, 14(9). Retrieved from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2560/2277>
- Hilton, J., Wiley, D., Stein, J., & Johnson, A. (2010). The four R’s of openness and ALMS Analysis: Frameworks for open educational resources. *Open Learning: The Journal of Open and Distance Learning*, 25(1), 37–44.
- Hylén, J. (2006). Open educational resources: Opportunities and challenges. *Proceedings of the 2006 Open Education Conference*. Retrieved from http://www.knowledgeall.com/files/Additional_Readings-Consolidated.pdf
- Koper, R. (2003). Combining reusable learning resources and services with pedagogical purposeful units of learning. In A. Littlejohn (Ed.), *Reusing online resources: A sustainable approach to e-learning* (pp. 46–59). London: Kogan Page.
- Lessig, L. (2004). *Free culture*. New York: Penguin Press. Retrieved from <http://free-culture.org/get-it/>
- McMartin, F. (2008). Open educational content: Transforming access to education. In T. Iiyoshi & M. S. V. Kumar (Eds.), *Opening up education* (pp. 135–148): Cambridge, MA: MIT Press.
- National Knowledge Commission. (2007). *Report to the nation 2007*. New Delhi, Government of India. Retrieved from <http://www.knowledgecommission.gov.in/download>

<loads/report2007/eng/Report07.pdf>

- OECD (Organisation for Economic Co-operation and Development). (2007). *Giving knowledge for free: The emergence of open educational resources*. Retrieved from <http://www.oecd.org/dataoecd/35/7/38654317.pdf>
- Petrides, L., Nguyen, L., Jimes, C., & Karaglani, A. (2008). Open educational resources: Inquiring into author use and reuse. *International Journal of Technology Enhanced Learning*, 1(1–2), 98–117.
- van der Baaren, J., & de Vried, F. (2010). How to foster sharing of educational resources? *Open Ed 2010 Proceedings*. Barcelona: UOC, OU, BYU. Retrieved from <http://hdl.handle.net/10609/4881>
- Vuorikari, R., & Koper, R. (2009). Evidence of cross-boundary use and reuse of digital educational resources. *International Journal of Emerging Technologies in Learning*, 4(4), 51–56. Retrieved from <http://online-journals.org/i-jet/article/view/1028>
- Wiley, D. (2009). *Dark matter, dark reuse, and the irrational zeal of a believer*. Retrieved from <http://opencontent.org/blog/archives/905>
- Windle, R., Wharrad, H., McCormick, D., Laverty, H., & Taylor, M. (2010). *Sharing and reuse in OER: Experiences gained from open reusable learning objects in health*. Retrieved from <http://jime.open.ac.uk/jime/article/viewArticle/2010-4/html>

Athabasca University 

