

Repositories post 2010: embracing heterogeneity in AWE, the Academic Working Environment

Peter Sefton
pt@ptsefton.com

Duncan Dickinson
duncan@dickinson.name
Central Queensland University

OpenRepositories July 2010, Madrid, Spain

Abstract: The organizers of the 5th International Conference on Open Repositories list nine polar dichotomies that represent “The Grand Integration Challenge” for the repository community/movement. In this paper we take up the challenge. We do so in the context of our work to build infrastructure for the academy in general, with the goal to develop a modular 'Academic Working Environment' (AWE) which encompasses both teaching and learning on one hand and research on the other. Repositories and the ecosystem of services and workflows that surround them play a key role in this emerging system.

1 Introduction

The “Grand Integration Challenge” set out for the 5th International Conference on Open Repositories 2010 (OR2010) is used in this paper as a focal point for discussing our work on the “Academic Work Environment” (AWE). In this challenge, the conference organisers laid out contemporary issues in repositories and asked how the academy can meet the challenge. Originating in our work for the Australian Digital Futures Institute at the University of Southern Queensland, the AWE is a set of activities oriented towards the research and development of integrated scholarly systems.

The Academic Working Environment is conceptualised as a loosely coupled set of services including the following components:

- A Learning Management System (LMS) / Virtual Learning Environment (VLE) and ePortfolio system for students and staff.
- A content management system and repository for courseware development and delivery.
- Repositories for research outputs including publications, creative works and datasets.
- A nascent Virtual Research Environment.
- Back-end, proprietary university systems for authentication, payroll, student management etc.
- “Personal” versions of the above – where participants in the academic environment have adopted their own preferred systems.

Through this paper we look at each of the OR2010 integration challenges listed on the call for papers, as listed in the sections below. We have bundled some of the related topics together not only for brevity but in recognition that, like the desktop and the cloud, the boundaries are increasingly blurred:

1. the web and the repository,
2. knowledge and technology,
3. wild and curated content,
4. linked and isolated data,
5. disciplinary and institutional systems,
6. scholars and service providers,
7. ad-hoc and long-term access,

8. ubiquitous and personalized environments,
9. the cloud and the desktop.

2 Meeting the grand challenges for repositories

The Academic Work Environment (AWE) is a working title for a set of services and computational systems that support the academic enterprise. The name was coined to capture under a single identifier a range of research and development work going on in a group tasked with pragmatic, practical work on workflows and computer systems for eLearning and eResearch.

2.1 The web and the repository & The cloud and the desktop

It is telling that *the web vs the repository* is posed as a dichotomy or a challenge. While repositories are, by default, web-based systems the vast majority of document content housed in repositories is in a non-web format, PDF. Furthermore, vast amounts of potential repository content such as data-files are excluded from the repository, and from the web/cloud because of the lack of services that assist academic users in making their content available on the web. The Academic Work Environment project works towards addressing both of the issues identified above with systems to:

1. Allow academic documents to be made available in HTML as well as PDF. This work, based on The Integrated Content Environment (ICE) was presented at Open Repositories 2009 (P. Sefton, Downing, & Day, 2009). Uptake on this has been very slow, but repositories need to start using web formats if they are to be part of the web and fulfill the promise of integrated documents and data as envisaged by Murray Rust and Rzepa (Murray-Rust & Rzepa, 2004).
2. Begin to close the gap between the desktop work environment and the repository via an application which brings a web-based repository view to all the files that a researcher/educator is using; allowing them to describe them, back them up and have them routed to appropriate repositories for works-in-progress and completed research outputs. We have explored this in a Virtual Research Environment application known as The Fascinator: Desktop edition. (Dickinson & P. Sefton, 2009a; P. M. Sefton, 2009)

There are a number of efforts in the repository space to create web-based software that operates 'upstream' of the institutional repository including:

- Rochester's IR+ (Tennant, 2009) which includes an area for researchers to do their work.
- Islandora offers workflow on top of the Fedora Commons repository platform (Leggott, 2009).
- RepoMMan provides local file browsing and repository upload (Richard Green & Chris Awre, 2009).
- Hydra works in a similar way¹ (C. Awre et al., 2009)

The systems above interact with the desktop only at the level of picking files. Some other applications have attempted to provide desktop services to label and sort data before uploading it to a repository. On the open source side, Field Helper, a tool for field workers such as archaeologists to categorize data and upload it to repositories was developed to a prototype stage at the University of Sydney, and is undergoing development after a hiatus, and Lensfield is being developed the University of Cambridge to allow physical chemists to manage data transformations prior to repository deposit (Downing, 2009). In the commercial arena, a product called Mediaflux is being used at some Australian institutions, for example to manage neuroscience data (Lohrey, 2007).

Other work in this area includes Microsoft Research's work on embedding SWORD – named OfficeSWORD² – repository deposit into their word processing product as well as some attempts to allow structured authoring and semantic web-authoring in the word processor (P. F. Fernicola, 2009; Fink et al., 2010) Another Microsoft Word Add-in embeds genetics workflows in Microsoft Word. None of these efforts so far offer any kind of generic mechanism for plugging in new kinds of semantics for new domains, or address the problems of interoperability with other word processing tools. Moving up from the desktop, Neylon (Neylon, 2009) looks at web-based processes for conducting research in a web context.

One fundamental issue with bringing the web and the repository closer together is resource packaging. One of the main features of PDF is that it packages a document into a single file, whereas web documents are a collection of resources which as far as the web architecture is concerned are unrelated. There are a number of different packaging formats that might be used in academia including plain compression formats such as Zip and various metadata formats like METS – for a discussion of the many options for describing packages of resources see the DRIVER II report (van Godtsenhoven et al., 2009). The profusion of potential formats for packaging reminds us of the old joke “the good thing about standards is there are so many of them”; with so many different protocols adopted in different systems workable interoperability is very hard to achieve.

1 See <http://projecthydra.org/>

2 <http://officesword.codeplex.com/>

One of the most promising of these packaging formats is the Open Access Initiative, Object Resource Exchange OAI-ORE, which (Lagoze et al., 2008) promised to remedy this – but it did not solve the problem of how to move around research publications as single files. ORE is not supported natively by web browsers or authoring tools, although there are a handful of plugins available for various tools.

Our experience with ORE – where we had to choose between several different software architectures for connecting services, none of them particularly user friendly (P. Sefton & Downing, 2010)– has led us to believe that what is needed to bring repositories and the web closer together is a resource packaging mechanism that is supported by mainstream tools; chiefly web browsers, but also authoring tools, content management systems and repositories. Two technologies that show promise in this area are:

1. EPUB – an HTML-based eBook format and packaging specification. EPUB has the advantage that it can contain simple HTML with images, but also richer content. EPUB files are single units, consisting of a Zip-compressed set of resources with a manifest (Wikipedia contributors, 2010a)
2. HTML 5 (Wikipedia contributors, 2010b) which introduces a manifest that can be used to bundle together a set of resources as a single entity or 'App'. This mechanism means that for the first time web users, including those using mobile devices will be able save documents into an interoperable bundle which could be emailed, or uploaded into a repository. Unfortunately, at this stage, this functionality seems to be limited to some mobile platforms, and does not work in desktop browsers.

It is not clear which of these would make the preferred 'master' format for scholarly materials; whether to use a rich HTML app with an embedded EPUB for simple reading, or the other way around; an EPUB document with a simple HTML view of the text of a document, but with additional richer resources available for readers that can support them. Both of these formats can include print-ready versions of a document in PDF format. Technically, creating them is not a major challenge, what will matter is the user experience and user preference, so we plan to test both.

A third option, the forthcoming PDF 2 specification could also be a contender as a container format for academic publications.

Another major issue is how to identify resources. The nature of the web, and the way resources are used offline as well as online makes it very difficult to assign persistent unique identifiers to particular resources – for examples of work in this area, see TripFS (Schandl & Popitsch, 2010) and the desktop URI scheme (Sauermaun, 2009). Schemes such as Handles (Sun, 2001) have gained some traction in the repository and publisher world (via DOIs) but our investigations have found no practical way to mint handles for resources as they are created on the desktop or in the lab and then be able to maintain a useful database of where things are as they move through academic workflows. Our group did implement persistent identifier support using Handles in the ReDBox Research Data management application (P. Sefton, Picasso, & Morgan, 2010), as this involved server infrastructure where governance can be put in place to maintain handle records properly, something which is not possible in a live research context at this stage.

There is a research and development question here around what should be packaged together, and what applications are best used for this. While we are convinced the word processor is still a key tool for academic writing, it is less clear whether it is also a suitable locus for embedded research workflows, particularly at the cost of interoperability.

2.2 Knowledge and technology

Knowledge and technology are terms that have at least three broad uses: in general English; as technical terms in disciplines such as knowledge management; and, sadly, as empty buzz-words. For the purposes of this brief discussion, we will focus on knowledge rather than technology. Knowledge, for the purposes of this discussion, is closely related to *understanding*; people (and possibly at some point machines) construct knowledge from interaction with information sources and with each other. Our specific interest in this is to reduce ambiguity in repository materials, mainly in the areas of metadata and document semantics – making it clear who wrote what and disambiguating technical terms.

Search engines provide a basic entry point to the information network but their reliance is on key phrases and not the context and meaning understood by a community of practice (Neumann & Prusak, 2007). The challenge now is to open up the information existing in online forms – including documents and data – so that it is findable³, a concept that spans more than just matching keywords:

Any system aiming to integrate heterogeneous data on an ad hoc basis and present this to users will need to adopt sophisticated models of relevance, quality, and trust that are sensitive to the user's current task and its context. (Heath Heath, 2008 p. 91).

Heath's discussion here is focused on the Semantic Web. Based on technologies from groups such as the W3C⁴, the Semantic Web provides a framework within the Web that gives web browsers and search engines the ability to interact with information.

3 <http://en.wikipedia.org/wiki/Findable>

Navigation in this model is based on assembling meaning rather than merely providing presentation services (Heath, 2008). The challenge, then, is to get the semantics into the Semantic Web. We are looking at methods that allow researchers to easily provide semantic information in their data and documents.

Specifically, we have explored ways that semantics can be embedded in academic documents, again focusing on desktop tools, mainly through word processors. Our approach has been to start with one of the simplest and most ubiquitous kinds of document semantics, metadata, asking how can we reliably and interoperably associate document metadata with text in-line, with the hope that the techniques can be expanded to deal with other semantics and relationships to research data. Initial work looked at encoding metadata in word processing documents using styles and tables (P. Sefton, Barnes, Ward, & Downing, 2009). More recently we have been working on more robust, simpler schemes for encoding formal semantic relations as URLs, a ubiquitous and widely supported technology. An early blog post⁵ describes how authorship can be asserted using a portmanteau URI⁶ such as:

```
http://ontologize.me/meta/?
tl_p=http://purl.org/dc/terms/creator&tl_o=http://trove.nla.gov.au/people/54165?
prop=dc:creator
```

Where the relationship (in RDF) is:

```
<referring page> <dc:creator> <http://trove.nla.gov.au/people/54165>
```

Moving from embedded metadata to a more distributed system for semantics, our Anotar⁷ project presented a framework for adding semantics through the well established concept of tagging and extending this by allowing taxonomies to be utilized. Anotar is explored in more detail below.

2.3 Wild and curated content

Our approach to wild and curated content is to gradually tame and domesticate the content, by allowing it to be husbanded by a series of 'curation events'. Using the Fascinator Desktop, the initial creator may label data items or sets in simple ad-hoc terms using tags such as "My Thesis" or "Anthropology 101 Course Notes". The incentive to do this is that when they do so, the items will be (a) backed up appropriately and (b) routed to collaborators automatically. This represents what we might call an emergent workflow; where object state changes result in items making progress through required stages.

We also provide a more intentional kind of curation via 'acts of publishing' where a data owner can push content across curation boundaries (the term coined by ARROW project members (Treloar, Groenewegen, & Harboe-Ree, 2007)) for various reasons;

- a draft can be pushed to a blog (remember that, using ICE content services, all word processing documents in the Academic Working Environment are web-ready and can be delivered as HTML)
- a finished paper or approved draft can be pushed to the institutional repository. Work is under way to enable this to happen in a way that (a) data is also deposited and (b) both document and data are web-ready and can be viewed in-browser as native web content, as well as being available for print.
- We are building systems for the Australian National Data Service to allow data to be registered with Research Data Australia.

2.4 Linked and isolated data & Ad-hoc and long-term access

Bootstrapping the linked-data web remains a grand challenge, but we are attempting to address it in work on The Fascinator Desktop by providing URIs (the formal name for links) for data while it is still in isolation in a lab or on a laptop computer for example. Managing identifiers through the lifecycle of a digital object is not easy when you consider that a researcher may have their digital files spread across multiple desktop, portable and mobile devices and that, in the messy landscape of desktop filesystems, filenames are changed at a whim and multiple versions may exist throughout a system.

Within The Fascinator, every item will have a URI from the moment it is discovered on the user's desktop. Creating a URI for desktop files is a similar approach to the SemDeskURI Scheme suggested by members of the Nepomuk⁸ project (Sauermaun, 2008) However, instead of relying on a new protocol (in the case of SemDeskUri this is "desktop://"), URIs in The Fascinator will utilise

4 <http://www.w3.org/2001/sw/>

5 <http://ptsefton.com/2010/11/14/before-beyond-the-pdf-authoring-tools-for-document-semantics.htm>

6 This paper uses a similar approach to identify its authors – the result is that our names are marked using RDFa semantics in the submitted version.

7 <http://www.purl.org/anotar/>

“http://”. For example, whilst “http://localhost/fred@example.com/research/data” will not allow remote users to access the resource, it does provide for contextual identification.

In terms of ad-hoc vs long-term access, under this scheme we would expect resources to move from an isolated desktop-web view of digital objects (remembering that our systems give people a web view from the very creation of the object) to ad-hoc team views where the provenance of items is preserved by keeping both desktop and team-URIs, through to more formally created views. As content is routed from the desktop to shared repositories the plan is to keep the URIs, so that the metadata for an item contains all of the known identifiers that we have for it.

2.5 Disciplinary and institutional systems / Scholars and service providers

Whilst much of this essay has focused on the technical aspects of “The Grand Integration Challenge”, the complexity of institutional and individual engagement must be considered. This is a challenge that threads through numerous areas of academia, including postgraduate researcher skills, central ICT provision, Government reporting requirements, Library systems, research project management etc. Indeed, it presents a “wicked problem”⁹ for the academic research sector and is one that various stakeholders approach with a carrot and/or stick approach.

From the carrot side, we work with motivated pilot users and faculties to build the AWE to meet their needs. Another carrot, of course, is grants - such as those offered by the Australian National Data Service – to teams developing solutions. The stick often presents itself as institutionalised mandates

(Australian Government, 2007; “Data-sharing culture has changed,” 2009)(Australian Government, 2007; “Data-sharing culture has changed,” 2009) but for many researchers, the technical complexities around data management are an intrusion into their time (Henty, Weaver, Bradbury, & Porter, 2008). As mentioned earlier, a central goal of the AWE is to provide services that meet the various stakeholder demands but don't interfere with the researcher's core tasks.

2.6 Ubiquitous and personalized environments

Two central goals of the Academic Working Environment, are:

1. to hide the technicalities around data management and the semantic web
2. and to provide services that meet various stakeholder demands.

These goals are both aimed at allowing the researcher to focus on their research. We envisage a mesh of repository services, building on the existing standards for linking repository content; the Fascinator Desktop work (and before it The Integrated Content Environment) introduced the idea of a personal desktop web; with services such as tagging and note-taking acting not only in their traditional role of assisting in the research process, but as triggers in emergent workflows.

A good example to illustrate how the personal and ubiquitous can meet and “intermesh” is our work on a general framework for annotations, Anotar. In work extending from that described in Dickinson and Sefton (2009), we added Anotar annotation services to The Fascinator. This work was motivated by a few requirements:

- Adding text, image and video and tagging annotation to a new web-based system, (The Fascinator) but wanting to do so in a way that user interface code could be reused on other systems in the Academic Working Environment such as Moodle.
- An easily re-deployable web-client framework to add such services.
- Something that is achievable now, but which will work with broader scholarly investigations of annotation.

We are aware of work that has been done in the Open Annotation Consortium on data modeling (Hunter, Cole, Robert Sanderson, & Van de Sompel, 2010), and will align the data-models and protocols used by Anotar as specifics emerge from that project.

8 <http://nepomuk.semanticdesktop.org/xwiki/bin/view/Main1/>

9 http://en.wikipedia.org/wiki/Wicked_problem

In the case of our work with USQ's Public Memory Research Centre, on a Vietnam War history project (Dickinson & P. Sefton, 2009) the annotation facility allowed the researcher to tag photos in their desktop repository for release to an online repository for viewing. Research participants would be able to log onto this repository and provide the researcher with essential information regarding the people and locations in various photos and movies through the use of taxonomy-based tagging. Furthermore, the open-ended nature of the annotations provide an online forum for the participants to share their memories and even debate points of view – providing a rich data-set for the researcher as well as a first-person public memory archive.

Whilst tagging and annotation services are provided (to some extent) by various online Web 2.0 systems, the AWE solution meets research-specific data management requirements by keeping research data in a way that adheres to the University's ethical clearance requirements. For the researcher and their community of participants, it provides a personalized environment that lets them focus on their own goals rather than the technical infrastructure.

3 Conclusion

The Academic Working Environment is not a single monolithic application with a “one size fits all” approach. It is a vision, or road map to an environment of interoperable services that works with the researcher and doesn't hamper their efforts.

We have touched on all the major themes of the OR2010 Grand Challenge in this review of work that has been conducted in our research group. We firmly believe that the AWE is not a fixed point but an evolving ecosystem and, in conclusion, we point the way to more work that is required to advance the themes discussed above:

- More effort is still required to mesh repository services with upstream working environments – this is a major challenge given the range of systems and environments used by researchers and other repository contributors, but as noted above, current projects should point the way to further work.
- One topic which is clearly 'hot' and ready to be addressed in 2011-12 is resource packaging, looking at EPUB and HTML 5 apps as a way to supplement or supplant PDF as a unit of academic currency.
- There is more work to be done to align the library-centric view of institutional repositories with their uses in other academic contexts, where the simple information models we have been using for publications need to be extended to deal with the variety of academic endeavours, ranging across open access learning materials, research data, and creative works.

4 References

Australian Government. (2007). *Australian code for the responsible conduct of research*. Canberra, Australia: Australian

Government. Retrieved from http://www.nhmrc.gov.au/_files_nhmrc/file/publications/synopses/r39.pdf

Awre, C., Cramer, T., Green, R., McRae, L., Sadler, B., Sigmon, T., Staples, T., et al. (2009). Project Hydra: Designing & Building

a Reusable Framework for Multipurpose, Multifunction, Multi-institutional Repository-Powered Solutions. Retrieved from <http://smartech.gatech.edu/handle/1853/28496>

Data-sharing culture has changed. (2009, November 12). *Research Information*. Retrieved November 19, 2009, from

http://www.researchinformation.info/news/news_story.php?news_id=553

Dickinson, D., & Sefton, P. (2009a). Creating an eResearch desktop for the Humanities. Presented at the eResearch Australasia

2009, Manly, Australia. Retrieved from <http://eprints.usq.edu.au/6090/>

Dickinson, D., & Sefton, P. (2009b). Creating an eResearch desktop for the Humanities. Presented at the eResearch Australasia

2009, Sydney. Retrieved from <http://eprints.usq.edu.au/6090/>

- Downing, J. (2009). lensfield - Google Code. *Project website*. Retrieved June 29, 2009, from <http://code.google.com/p/lensfield/>
- Fernicola, P. F. (2009). Incorporating Semantics and Metadata as Part of the Article Authoring Process. Retrieved March 1, 2010, from http://elpub.scix.net/cgi-bin/works/Show?152_elpub2009
- Fink, J. L., Fernicola, P., Chandran, R., Parastitidas, S., Wade, A., Naim, O., Quinn, G., et al. (2010). Word add-in for ontology recognition: semantic enrichment of scientific literature. *BMC Bioinformatics*, *11*(1), 103. doi:10.1186/1471-2105-11-103
- van Godtsenhoven, K., Elbaek, M. K., Sierman, B., Bijsterbosch, M., Hochstenbach, P., Russell, R., & Vanderfeesten, M. (2009). Emerging Standards for Enhanced Publications and Repository Technology: Survey on Technology. Retrieved from <http://dare.uva.nl/aup/nl/record/316870>
- Green, Richard, & Awre, Chris. (2009). Towards a Repository-enabled Scholar's Workbench. *D-Lib Magazine*, *15*(5/6). doi:10.1045/may2009-green
- Heath, T. (2008). How Will We Interact with the Web of Data? *Internet Computing, IEEE*, *12*(5), 88-91.
- Henty, M., Weaver, B., Bradbury, S. J., & Porter, S. (2008). *Investigating Data Management Practices in Australian Universities*. Retrieved from <http://eprints.qut.edu.au/14549/>
- Hunter, J., Cole, T., Sanderson, Robert, & Van de Sompel, H. (2010). The Open Annotation Collaboration: A Data Model to Support Sharing and Interoperability of Scholarly Annotations. *Digital Humanities 2010* (pp. 175-177). Retrieved from <http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/pdf/book-final.pdf#page=201>
- Lagoze, C., Van de Sompel, H., Nelson, M. L., Warner, S., Sanderson, R., & Johnston, P. (2008). Object re-use & exchange: A resource-centric approach. *Arxiv preprint arXiv:0804.2273*. Retrieved from <http://arxiv.org/abs/0804.2273>
- Leggott, M. A. (2009). Islandora: a Drupal/Fedora Repository System. Proceedings, . Retrieved November 30, 2010, from <http://smartech.gatech.edu/handle/1853/28495>
- Lohrey, J. (2007). Mediaflux: a data management platform for collaborative research.
- Murray-Rust, P., & Rzepa, H. S. (2004). The Next Big Thing: From Hypermedia to Datuments. *Journal of Digital Information*, *5*(1), 248.
- Neumann, E., & Prusak, L. (2007). Knowledge networks in the age of the Semantic Web. *Briefings in Bioinformatics*, *8*(3), 141-149. doi:10.1093/bib/bbm013
- Neylon, C. (2009). Head in the clouds: Re-imagining the experimental laboratory record for the web-based networked world. *Automated Experimentation*, PubMed Central, *1*, 3-3. doi:10.1186/1759-4499-1-3

- Sauermann, L. (2008, October 27). RFC-draft: SemDesk URI Scheme. Retrieved February 25, 2010, from http://dev.nepomuk.semanticdesktop.org/repos/trunk/doc/2008_09_semdeskurische/index.html
- Sauermann, L. (2009, July 30). SemdeskUris – aperture. Retrieved June 23, 2011, from <http://sourceforge.net/apps/trac/aperture/wiki/SemdeskUris>
- Schandl, B., & Popitsch, N. (2010). Lifting File Systems into the Linked Data Cloud with TripFS. Presented at the 3rd International Workshop on Linked Data on the Web, Raleigh, North Carolina, USA. Retrieved from <http://eprints.cs.univie.ac.at/69/>
- Sefton, P. M. (2009). The Fascinator - Desktop eResearch and Flexible Portals. Presented at the 4th International Conference on Open Repositories, Georgia, U.S.A.: Georgia Institute of Technology. Retrieved from <http://smartech.gatech.edu/handle/1853/28483>
- Sefton, P., & Downing, J. (2010). ICE-Theorem - End to end semantically aware eResearch infrastructure for theses. *Journal of Digital Information*, 11(1). Retrieved from <http://journals.tdl.org/jodi/article/viewArticle/754>
- Sefton, P., Barnes, I., Ward, R., & Downing, J. (2009). Embedding Metadata and Other Semantics in Word Processing Documents. *International Journal of Digital Curation*, 4(2). Retrieved from <http://www.ijdc.net/index.php/ijdc/article/view/121>
- Sefton, P., Downing, J., & Day, N. (2009). ICE-theorem - end to end semantically aware eResearch infrastructure for theses. *University of Southern Queensland*. Retrieved August 24, 2009, from <http://eprints.usq.edu.au/5248/1/ice-theorem-paper-OR09.htm>
- Sefton, P., Picasso, V., & Morgan, T. (2010). Balancing business imperatives and leveraging capability: a model for research data management. Presented at the eResearch Australasia 2010, Gold Coast, Queensland, Australia. Retrieved from <https://ocs.arcs.org.au/index.php/eraust/2010/paper/viewPaper/79>
- Sun, S. (2001). Establishing persistent identity using the handle system. *Proceedings of the Tenth International World Wide Web Conference*. Presented at the Tenth International World Wide Web Conference.
- Tennant, R. (2009). Rochester Releases Their “IR+” Repository Platform « Tennant: Digital Libraries. Retrieved November 30, 2010, from <http://blog.libraryjournal.com/tennantdigitallibraries/2009/12/16/rochester-releases-their-ir-repository-platform/>
- Treloar, A., Groenewegen, D., & Harboe-Ree, C. (2007). The Data Curation Continuum. *D-Lib Magazine*, 13(9/10). doi:10.1045/september2007-treloar
- Wikipedia contributors. (2010a, November 28). EPUB. *Wikipedia, The Free Encyclopedia*. Wikimedia Foundation. Retrieved from

<http://en.wikipedia.org/w/index.php?title=EPUB&oldid=399335606>

Wikipedia contributors. (2010b, November 30). HTML5. *Wikipedia, The Free Encyclopedia*. Wikimedia Foundation. Retrieved from <http://en.wikipedia.org/w/index.php?title=HTML5&oldid=399647428>

Copyright Peter Sefton and Duncan Dickinson, 2010. Licensed under Creative Commons Attribution-Share Alike 2.5 Australia.

<http://creativecommons.org/licenses/by-sa/2.5/au/>

