

Development of a Publishing Framework for Living Open Access Textbooks

Anita EPPELIN^{A,1}, Richarda BÖTTCHER^B

^a*German National Library of Medicine (ZB MED), Cologne, Germany*

^b*Trauma Hospital Berlin (UKB), Berlin, Germany*

Abstract. We are suggesting a framework for open access publishing of comprehensive, elaborately structured digital textbooks, potentially enriched by extensive non-textual data. The framework will comprise of a software toolbox containing a collaborative authoring platform, manuscript workflow system, tools for editorial work, presentation platform, updating workflow, interfaces for dissemination and long-term preservation. In addition to the software toolbox itself, guidelines and standards for all aspects of digital textbook publishing, including peer review, metadata and editorial procedures, legal issues as well as a business model will be developed, resulting in a best practice guide to digital textbook publishing. The software and workflows will not be developed from scratch, but build upon existing open-source software, using the example of a scientific textbook of hand surgery as its first use case. This paper introduces the project and gives a prospectus of the proposed framework.

Keywords. open access, publishing, open educational resources, textbooks, medicine, hand surgery

Introduction

Unlike for journal e-publishing, there exist few, if any, established software tools and workflows for digital publishing of scientific textbooks. Customary content management software is relatively easy to implement, but offers only basic web publishing functionalities and therefore requires extensive customizing for fulfillment of the specific requirements of scientific publications. Existing publishing platforms developed for e-books or monographs currently offer a simple linear presentation, thereby imitating the print model and plainly transferring it to the online medium. While this approach might be sufficient for the demands of certain formats and certain disciplines like the humanities, it fails to exploit the full potential of online publishing for more complex structured publications like those in the STM area. Journal publishing evolves rather dynamically, with novel approaches addressing different stages of publishing. Examples are public peer review [1], policies for research data publishing [2], semantic enrichment of articles [3], and alternative impact metrics on article-level, taking into account new ways of dissemination, e.g. through social networking [4, 5]. However, no groundbreaking approaches for textbook publishing are

¹ Corresponding author Anita Eppelina, German National Library of Medicine (ZB MED), Cologne, Germany; Email: eppelin@zbmed.de.

known to the authors as yet (even in the STM area, where new technologies usually are adopted rather quickly).

In principle, the directions of future development by software support can be divided in three categories: (a) enhancement of content presentation (e.g. enabling annotation by the reader through comments or bookmarks, semantic structuring and further enrichment), (b) support of production process prior publication (e.g. support of peer review process, collaborative editing of documents, and seamless integration of other digital tools of the scientific process like wikis and “virtual research environments”), and (c) support of communication and collaboration, i.e. sharing, within the community (e.g. through integration of social networking features, commenting and updating workflows).

Several initiatives have emerged lately, promisingly addressing one or more of these three categories. For example, “Beyond the PDF”, a visionary theoretical concept mainly focusing on content presentation (a) on a rather technical level, has been developed [6]. A more practical approach of (a), Annotum, is based on the software WordPress, a content management software originally intended for blogs [7]. It combines the ease of use and popularity of WordPress with scientific publishing features, especially such that are applicable to journal publishing, in particular several dissemination interfaces. In terms of supporting the process prior publication (b), one example is Open Monograph Press, an “online workspace for publishing monographs” based on the same software like the well-known Open Journal Systems by the Public Knowledge Project [8]. Another example for (b) is Textus, mainly intended for the humanities [9]. Under the heading “Open Educational Resources” (OER), a large-scale project for undergraduate education textbooks is in an advanced stage, with strong focus on the sharing aspect (c) [10].²

An integrated approach of supporting tools for all of the mentioned directions is still lacking. Certainly, exploiting the technical possibilities aiming at a fully integrated software support platform should be questionable particularly in the agile field of open access publishing and open source where a bespoke solution will much easier be accepted and adopted by the community than any standard solution ignoring the special community needs. For example, a standard student’s textbook might need an environment for the production process but will hardly ever be topic of community discussion. A journal paper written by one or two authors might get a lot of community attention and comments but is probably written even without any versioning tool. Both might need some software support, but an integrated toolbox is not necessary.

However, the mentioned examples represent just a part of the scientific communication. Many special disciplines use scientific textbooks as important part of as possible up-to-date knowledge. Thus, scientific textbooks – read by scientists and advanced students as well – have two tasks: containing the entire detailed knowledge of the field, being updated as soon as any new scientific advance is reached. These in some manner diverging aims of collecting and updating are in the focus of an integrated publishing toolbox.

The project presented here will address this demand by developing a framework for open access publishing of “living” scientific textbooks. It will re-use existing open-source software and build upon expertise in e-publishing of scientific contents as far as possible.

² A comprehensive catalogue of publishing software can be found under <http://www.carpet-project.net/en/>.

The toolbox shall encourage the entire writing and publishing process including review and updating procedures providing standards for versioning and archiving. It shall integrate social network features for dissemination, but also as source of new textbook updates. Furthermore, it shall enable the reader to work with the textbook as a real online document getting access to further additional material (e.g. raw data, movies) that could not have been added to an ordinary book.

1. Project Outline

1.1. Aim of the Project

First, the project aims at developing a state-of-the-art software toolbox for open access publishing of digital textbooks. It will be able to deal with very comprehensive, elaborately structured contents, potentially enriched by extensive non-textual data material like graphics, multimedia, visualizations of raw data and many others. Moreover, it shall offer various features, in particular for widest-possible dissemination of the contents. By a requirements analysis, it is ensured that the software functionalities will be able to meet a wide range of demands by the scientific community. In general, it is intended to develop a generic, re-usable and customizable infrastructure that bears the potential to be suitable for all scientific disciplines.

Second, the project aims at gaining experience on how to use the given technical options to full capacity. The overall aim is to enable the scientific community to establish enhanced, faster and more transparent communication practices. This part of the project will initially be based on the use case described below, with the potential of being extended by experiences from similar publishing projects. It will result in a collection of best practice recommendations for all aspects of the publishing life cycle, e.g. for reviewing, commenting or updating of contents.

1.2. Frame Conditions, Sustainability and Open Access Principles

The project has been initiated and is carried out by the German National Library of Medicine (ZB MED), a publicly funded institution for scientific information infrastructure. One of ZB MED's strategic goals is to offer scientists support during the whole research process, i.e. not only in terms of literature supply. Communication of research results and experiential knowledge in terms of education are crucial for scientific progress and one of the genuine motivations of all scientific activities. Therefore, ZB MED develops and maintains publishing services designed for the needs of scientists. One example is German Medical Science, a publishing platform for medical journals and conference proceedings running successfully since 2003 [11]. Thus, ZB MED is well prepared for the project. It is already employing staff with know how and experiences in management of publishing workflows and, moreover, maintains an environment of infrastructures ready to embed the project in (e.g. long-term preservation via Goportis [12], persistent identification via DataCite [13]).

After completion, ZB MED will continue to operate the system on a permanent basis as a service for the scientific community. It will seek to cooperate with professional associations in the field of life sciences and other partners with the aim to publish more textbooks like the use case described below. In addition, in order to enable scientists to self-publish their works independently from institutions like ZB

MED and other third parties, the completed toolbox will be made available as open source software for re-use and further development. The accompanying best practice guides will be published open access. The project will be carried out under the following core principles:

- Re-use of existing solutions as far as possible (mainly software, but also e.g. templates for legal texts like licenses or author's contracts)
- Transparent development process, i.e. early consecutive releases of pre-final (beta) stages of software and other relevant intermediary development products
- Open source distribution of resulting software products
- Open Access publishing of documentations, best practice guides, and all other resulting publications, if relevant
- Open Access publishing of all prospective contents to be published with the toolbox (including research data)
- Close collaboration with user groups throughout the project

1.3. Development Plan

The project is planned in four parts: pre-development, development, implementation of use case, and documentation. The following outline lists the different steps:

0. Pre-development phase

- a. Analysis of potential and requirements (central questions: How do scientists work when creating or consuming textbook contents, what kind of support do they need in doing so, and how might this change in the future?)
- b. Analysis of existing solutions and software products, in particular with regard to re-usability
- c. SWOT analysis
- d. Compilation of requirements specification
- e. Estimation of effort, feasibility evaluation

1. Development phase

- a. Development of software according to requirements specification, main components:
 - i. Collaborative authoring platform
 - ii. Manuscript submission and peer review workflow system
 - iii. tools for editorial work (text capture, markup, copyediting, proofreading)
 - iv. Updating and commenting workflow system
 - v. Interfaces for dissemination and long-term preservation
- b. Development of a set of principles and recommendations (best practice guide) for the publishing process, main components:
 - i. Peer-review standards and criteria
 - ii. Editorial principles and metadata standards
 - iii. Legal guidelines and contract templates
 - iv. Business model options
 - v. Standards for updating and commenting of contents

2. Phase of implementation of prototypical use case

3. Documentation phase

- a. Documentation and distribution of software
- b. Publication of best practice guide
- c. Documentation of project experiences (presentation at conferences, publications)

1.4. Functionalities and Specifications of the Proposed Framework

The toolbox will comprise of several components. First, there will be a collaboration and communication platform for authors and the editorial team. It will offer a workflow system for submission and peer review of manuscripts, images, videos and all other materials to be published. There will be tools for editorial work like text capture, structuring, formatting, copy-editing and proofreading. The heart of the toolbox will be a presentation platform, i.e. the textbook itself. A concurrent context navigation (including parts not published yet) will guide the user through the structure. Through fine-tuned editorial structuring, generation of different indices will be possible, like keyword index, acronym index, and also subject-specific indices, like e.g. an index of surgical techniques. The technology to be used will allow for a device-independent and barrier-free use of the contents. An ongoing discussion of the contents will be supported by enabling comments and annotations. Usage statistics will be gathered, also from external sources including social networking services, in order to reflect the impact of the publication according to the “altmetrics” principles [4] (at least as far as these are applicable in a meaningful way for textbook publications and their building blocks). For dissemination and collaboration, social networking features like social bookmarking will be integrated, in addition to more conventional interfaces like OAI-PMH aimed at more technically elaborate uses from service providers. There will also be an updating workflow (hence “living textbook”), as well as a versioning mechanism in order to ensure transparency of the documents’ history and the updating process. Published data will be made referable and thus citeable by a persistent identifier (DOI) and securely stored, and there will be an interface to ingest contents into a long-term preservation system.

For each publishing project, in addition to software development, there are a lot of issues that need clarifying. These include legal issues, peer review criteria, standards for metadata and editorial procedures, and, last but not least, a business model. Thus, the toolbox will be accompanied by a comprehensive best practice guide for open access publishing of scientific textbook (see also section 1.1. and 1.3.).

2. Case Study

The developmental and prototypical test case for the system will be a scientific textbook of hand surgery. Such a publication does not exist worldwide currently. However, there is a big demand for digital and, even more so, open access availability of such contents. Throughout their career, surgeons in particular refer to extensive textbooks in their respective fields of expertise rather than journal articles. Hand surgery is based on very specialized surgical education, research and discussions. In some fields diagnostic and therapeutic aspects are changing rather fast. In particular, surgical techniques can be imparted perfectly through audiovisual material. According to the principle of open access, everybody with internet access will get the possibility

to inform themselves about all aspects of hand surgical treatments, to view operation videos or to compare his own results with the state of the art.

The textbook will be published in English and is intended primarily for hand surgeons in education, but also for daily surgical practice. It will include principles and basics, clinical aspects and all information about diagnostics, treatment, and rehabilitation. The textbook will be published in several stages to completion and then updated continuously. An international editorial board of 10+ hand surgeons has been compiled. On the one hand, its members will be responsible for ensuring the scientific quality of the contents. On the other hand, they will also supervise and test the suitability of the workflows and toolbox to be developed.

As stated above, the whole process of creation of the prototype textbook will be documented and be integrated in a best practice guide for open access textbook publishing.

3. Conclusion and Outlook

The aim of the project is to develop an innovative, yet sustainable publishing toolbox for scientific textbooks. By close collaboration with the scientific community, it strives for fulfilling the demands of all involved parties – editors, authors, reviewers, and users – the best-possible way. According to the open access principle, the toolbox with its applications and documentations as well as the accompanying best practice guide will be freely available for access as well as re-use.

Given these possibilities, scientific textbooks could also effectively profit of the vast every-day experience of users with more popular internet community projects like Wikipedia. Generally speaking, the framework proposed here might be a building block in order to enable the scientific community to create and establish new forms of documenting their knowledge.

References

- [1] Pöschl U, Koop T. Interactive open access publishing and collaborative peer review for improved scientific communication and quality assurance. *Information Services & Use* 2008;28:105-7. <http://dx.doi.org/10.3233/ISU-2008-0567>
- [2] Piwowar H, Chapman WW. A review of journal policies for sharing research data. *ELPUB* 2008. <http://ocs.library.utoronto.ca/index.php/Elpub/2008/paper/view/684> (last access: March 26, 2012)
- [3] Shotton D, Portwin K, Klyne G, Miles A. Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article. *PLoS Comput Biol* 2009;5(4):e1000361. <http://dx.doi.org/10.1371/journal.pcbi.1000361/>
- [4] Neylon C, Wu S. Article-Level Metrics and the Evolution of Scientific Impact. *PLoS Biol* 2009;7(11):e1000242. <http://dx.doi.org/10.1371/journal.pbio.1000242/>
- [5] Priem J, Piwowar HA, Hemminger BA. Altmetrics in the wild: Using social media to explore scholarly impact. 2012. <http://arxiv.org/abs/1203.4745v1> (last access: March 26, 2012)
- [6] Beyond the PDF. <http://sites.google.com/site/beyondthepdf/>, 2011 (last access: March 26, 2012)
- [7] Annotum. <http://annotum.org>, 2012 (last access: March 26, 2012)
- [8] Willinsky J. Toward the Design of an Open Monograph Press. *Journal of Electronic Publishing* 2009;12(1). <http://dx.doi.org/10.3998/3336451.0012.103>
- [9] Gray J. TEXTUS: an open source platform for working with collections of texts and metadata. December 2011. <http://jonathangray.org/2011/12/08/textus-an-open-source-platform-for-working-with-collections-of-texts-and-metadata/> (last access: March 26, 2012)
- [10] Organisation for Economic Co-operation and Development (OECD). Giving Knowledge for free: The Emergence of Open Educational Resources. 2007. ISBN 92-64-03174-X

- [11] German Medical Science. <http://www.egms.de/>, 2011 (last access: March 26, 2012)
- [12] Lindlar M. Digitization and digital preservation projects at the German National Library of Medicine (ZB MED). JEAHIL 2010;6(2):7-9.
- [13] Brase J. DataCite revisited – Citing data in the XX1st century – at long last. Academic Publishing in Europe http://ape2012.eu/ppt_wednesday/15brase_ape2012.pdf (last access: March 26, 2012)