

The Public Knowledge Project: Open Source Tools for Open Access to Scholarly Communication

James MacGregor, Kevin Stranack and John Willinsky

Abstract This chapter describes how the Public Knowledge Project, a collective of academics, librarians, and technical geniuses, has been, since 1998, building open source software (free) publishing platforms that create an alternative path to commercial and subscription-based routes to scholarly communication. It sets out how its various website platforms, including Open Journal Systems, Open Conference Systems, and, recently, Open Monograph Press, provide a guided path through the editorial workflow of submission, review, editing, publishing and indexing. Thousands of faculty members around the world are now using the software to publish independent journals on a peer-reviewed and Open Access basis, greatly increasing the public and global contribution of research and scholarship.

Introduction

The digital transformation of scholarly communication has given rise to a wealth of new publishing strategies, models, and tools over the last two decades. Many of these developments revolve around this new technology's seeming promise to increase the extent and reach of knowledge dissemination on a more equitable and global scale (cf. Benkler 2006). At first reluctantly, but now with increasing levels of interest, scholarly publishing is turning to the Internet as the preferred method of dissemination. This poses a set of core challenges: How can scholars, regardless of geographic location or institutional resources, participate in, learn from, and contribute to the global exchange of knowledge? How can those involved in

J. MacGregor (✉) · K. Stranack
Simon Fraser University Library, Burnaby, Canada
e-mail: jmacgreg@gmail.com

J. Willinsky
Stanford University, Stanford, USA

publishing scholarly work maintain high standards of quality for this work, while further advancing the long-standing research goals of openness and increased access to knowledge?

The responses to such challenges are prolific and growing, proving to be source of exciting research and development for scholarly communication. The whole arena is complicated by the mix of, on the one hand, large corporate publishing entities, seeking to preserve some of the highest profit margins in publishing, and, on the other, small independent scholar-publishers who see their work as a service to their colleagues and a commitment to the pursuit of learning, while in the middle are scholarly associations that have grown dependent over the years on publishing revenues for their survival (cf. Willinsky 2009).

Within this larger picture, the Public Knowledge Project (PKP) represents a modest instance of research and experimentation in a new generation of publishing tools that would lower the barriers among scholars and scientists interested in playing a more active role in publishing and in seeing this work reaching a much wider audience. What success the project has achieved over the years in developing software that is used by journals in many different parts of the world can be attributed to the collective wisdom, as well as trial and error, of the team involved in this project. This wisdom has found its expression in, for example, the early adoption of open source and community development models; the active development of the international PKP community; and the feedback of users in guiding software and workflow design decisions that reflected principles of simplicity, interoperability, accessibility, and openness, without sacrificing capability.

History of the Project

The Public Knowledge Project was started at the University of British Columbia (UBC) in Vancouver, Canada in 1998 with a small number of student developers working under the direction of John Willinsky, a faculty member in the Faculty of Education. The project began on two levels. It was involved in researching various models for creating a more coordinated approach to scholarly publishing that would increase its public value (cf. Willinsky 1999). It also sought to develop something practical and useful, in the form of online software, for the journal community, which was only beginning, at the turn of the twenty-first century, to think about moving its publishing operations and its journals to a web-based environment. The project's goal from the outset was to find ways of increasing public and global access to research and scholarship. As such, it was an early participant in the Open Access movement that sought to develop ways of creating peer-reviewed journals that did not charge readers for access to their content (cf. Suber 2012).

In 2005, Willinsky led PKP into a partnership with Lynn Copeland and Brian Owen at the Simon Fraser University (SFU) Library. It was becoming clear that the research library was emerging as a key player in efforts to reform scholarly

communication around technology. And SFU Library quickly demonstrated how productively this sort of partnership could work, by providing management and systems support. In 2007, Willinsky left UBC to take up a position at Stanford University in the Graduate School of Education, creating a strong institutional link for the project between Stanford and the SFU Library, with a greater emphasis at Stanford on the related research questions in scholarly communications, while matters of technology development and innovation are centered around the SFU Library.

As for PKP's software, an early result of the project was the creation and 2001 release of Open Journal Systems (OJS) 1.0, a free, open source journal publication management system, which provided an online platform for accepting submissions, performing double blind peer review, editing, publishing, and dissemination. OJS provided the necessary technological infrastructure for many journals making the transition from print to online, as well as being the foundation for emerging "born-digital" journals. Today, PKP has not only created OJS, but also Open Conference Systems (OCS, for managing conferences), Open Monograph Press (OMP, for managing book publication), and Open Harvester Systems (OHS, for metadata harvesting).

The Project currently employs over a dozen developers, librarians and library staff at Simon Fraser University, Stanford University, and elsewhere in the world. Most are active academically as students and/or as scholars. There is no physical PKP "office": while most PKP associates live in Vancouver, others live in New Brunswick, Palo Alto, Brazil, and elsewhere. As with other open source software initiatives, community collaboration is at the forefront of the PKP development model, and the PKP user community continues to grow and influence application development.

To take one community example, the PKP Support Forum has over 4,500 members and more than 33,000 posts, and generates upwards of ten new posts a day; many of these posts contain vital bug reports and feature requests.¹

Translation support across all PKP applications is growing: OJS alone now includes 27 community-contributed translations. And the global install base of PKP applications expands from month to month, with over 3,500 active journals currently using OJS. Half of these titles are edited and published in developing countries. Close to 90 percent of the titles publish Open Access editions, with very few charging article processing fees. Their secret is operating on very low-overhead and obtaining some measure of institutional support (cf. Edgar and Willinsky 2010).

In both its research and development initiatives, PKP continues to be funded, as it was originally, by a range of different government and foundation grants.² In 2012, however, PKP introduced two new funding models to ensure its sustainability, which involved growing responsibilities around the expanded number of

¹ PKP: <http://pkp.sfu.ca/support/forum>

² PKP Funding Partners: http://pkp.sfu.ca/funding_partners

journals and conferences dependent on its software. The two models involved, first of all, strengthening PKP's hosting services for journals and conferences using its free software, and, secondly, the creation of an institutional sponsorship program for research libraries, many of which were now providing PKP software to their institutions.

On the hosting side, what was once a fledgling, ad-hoc initiative at Simon Fraser University Library to mount journals was established as a distinct venture, dubbed PKP Publishing Services (PKPIPS).³

The growth and professionalization of PKPIPS has required a deeper level of commitment to infrastructure: hardware, network uptime, and software management across hundreds of installed instances of OJS, OCS and OMP. PKPIPS currently hosts over 200 journals and conferences (with a small number of OMP instances on the way), and now acts as a significant funding resource for PKP, not to mention a critical vector for feedback from invested, informed, and day-to-day users of the software.

Perhaps more significant for giving PKP a stronger institutional base, however, is its sponsorship program, which has now over 30 participating institutions.⁴

Interested research libraries can sponsor the project directly on an annual basis, or can become more involved as development partners. Development partners are just that: they have access to the core PKP development team and are deeply involved in long-term technical and administrative planning. This represents a new model for PKP, which has traditionally been a very small and tight-knit group of developers. Opening the team to a larger community is not without its challenges in coordinating the work among different teams and locations. It is important to stress here, however, that this isn't simply a solution to PKP's financial problem. The sponsorship program provides a venue for PKP to interact with the larger scholarly community in a way that previously did not exist. It is an open invitation to participate as a patron and a peer in this project, and the investment of participation is equally if not more important to the fundamental goals of the project as any financial contribution.

The PKP Systems

In introducing the operating principles at work in PKP's software, we are focusing on the themes of simplicity and interoperability that underlie the approach to supporting the managers, editors, authors, reviewers, copyeditors, layout designers, and proofreaders, among others involved in the workflow that defines scholarly publishing. Our goal has always been to build systems that are not simply free to use, but are easier to use to do the quality work that those in scholarly

³ PKP Publishing Services: <http://pkpservices.sfu.ca>

⁴ PKP Sponsorships: <http://pkp.sfu.ca/sponsorships>

publishing has always involved. We have sought to build systems that not only support the workflow that results in a sharing of research but that are instructive and informative around the standards that have historically developed around these practices, so that others who have not been previously part of the scholarly publishing community could begin to participate and contribute, as they were walked through the process by the software design. We have, in this process, pursued a number of holy grails, among them, the design of intuitive systems and the production of automated systems. We continue down this path, not without our Monty Python moments, having realized that scholarly publishing is not an inherently intuitive process nor one that can be readily automated. We have reduced the clerical tasks and greatly increased the portability of the editorial office, and a good deal more than that, of course, as we outline in what follows.

(a) **Simplicity in Technical Administration**

All PKP application system requirements are both *low* and *broad*: all that is needed to run OJS, OCS or OMP is a web server running PHP and a common database system (MySQL or PostgreSQL). PKP also actively supports older versions of PHP, MySQL and PostgreSQL, out of consideration for users who may not have access to newer technology. Users who download and install PKP applications are often definitively non-technical, and so special care has been taken to ensure that the installation and maintenance processes and documentation is easy to understand and uncomplicated. The installation process is straightforward and can be accomplished in a matter of minutes. Application maintenance, including backing up files and general upkeep, is also simple and well documented. After installing OJS, OCS or OMP, the site administrator can create one or more journal, conference or press instance on the site. Each instance takes only a second to create; after they have been created, separate journal managers can be enrolled in each instance, and these journal managers subsequently take over day-to-day administration tasks.

(b) **Simplicity in Management**

After the journal, conference or press has been created, the manager completes a guided setup process where all core components related to publishing workflow, author and submission management, guidelines (author, reviewer, editing, etc.), publishing, indexing, and the look and feel of the site are configured. (In OJS, the setup is a five-step process; in OCS, it is a six-step process; in OMP the process is a bit more extensive, with separate setup steps for press, website, publication, distribution and user management; there is an initial wizard available for quick press configuration, however.)

This stepwise workflow has been created and adhered to with different goals in mind. For the new manager, these workflows provide a guided tour through many of the options they must consider before they publish: OJS includes prompts for ISSN information and a publication schedule, for example, while OMP provides comprehensive series and category configuration options. For the experienced

manager, these setup options are easily and centrally accessible for periodic review.

This isn't a case of "simple is as simple does," however. A great deal of behind-the-scenes automation and task/service management is included in OCS, OJS and OMP, and all three applications offer far more capability than may be assumed from their relatively straightforward configuration processes. Most of these services involve promoting accessibility and visibility of the journal's published content on the web. For example, Google Scholar requires article information to be available to its web crawlers in very specific ways; OJS does this automatically, with no further configuration needed.⁵

(c) **Simplicity of Submission**

Each application's submission process has been refined to be as simple as possible for new and experienced authors alike. Each application uses a multi-step submission process (no more than five steps in any application). Each step serves a specific purpose, from informing the author of any copyright or other requirements; to providing submission indexing metadata; to requesting submission and/or supplementary files; to confirming the submission. Authors are aware at all times of which step they are on, and what is needed of them. This process ensures that all information relevant to the submission is gathered at the very beginning, saving editors valuable time during later stages.

(d) **Simplicity of Review**

While implementations differ as required by the publishing format, all three applications approach review and editing workflows in a philosophically similar way. Peer review is the key quality control for scholarly communication, as well as a source of improvement for this work. The review process, in particular for reviewers, must be kept as simple and quick as possible, as reviewers often have the least incentive to use the system and may balk at any impediment between themselves and the review proper. Typically, in the review process, the reviewer needs to agree to complete the review; download the submission files; and upload review comments and/or review files to the system. Reviewers may log in directly to the system to complete the review process, or editors may act on their behalf. To assist editors in selecting reviewers, the system tracks a reviewers previous record on areas of interest, time taken, number of reviews, and editor rating.

(e) **Simplicity of Editing and Production**

All three systems handle editing differently. OCS includes a relatively minor editing step only if full paper submissions are being accepted by the conference, whereas OJS and OMP both have full-scale editing workflows which can include input from copyeditors, proofreaders, layout editors, and others. In the case of OMP, editing and production workflows are handled separately: copyediting of

⁵ See Google Scholar: <http://scholar.google.ca/intl/en/scholar/inclusion.html>

final draft files are handled in an editing stage, and the creation of production-ready files (eBooks, PDFs, and so on) and the completion of all catalog information for that particular manuscript are managed in a final production stage.

The PKP Program

(a) Enhancing Interoperability

The PKP development team actively pursues software interoperability with different applications, frameworks and platforms where appropriate. Interoperability is ideally facilitated via open, widely used and time-tested APIs, standards and protocols. A number of interoperability standards common to the library and scholarly publishing worlds have enjoyed a long history of support within PKP, and support for new standards is added regularly (and in many cases in the form of contributed plugins from the larger community).

Various information interchange mechanisms enjoy broad support across the PKP applications. All three applications support the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), which provides machine access to published article (or book, or presentation) metadata for the use of indexing systems. Another source of interoperability comes from following XML standards, particularly journal publishing standards such as the National Library of Medicine (NLM) Journal Publishing Tag Set, have proven crucial to PKP's efforts to provide open, structured access to published scholarly content. XML is particularly well-suited to sharing data and metadata online between applications, as it is human- and machine-readable.

Other discrete interoperability projects are currently underway. PKP is partnering with the Dataverse Network initiative at Harvard to develop a set of plugins that will provide deposit and display functionality between OJS and Dataverse repositories.⁶ At the same time, the project is also working with the Public Library of Science (PLOS) to provide Altmetrics⁷ for PKP applications, and with ORCID to provide author disambiguation services.⁸ These services are usually implemented as plugins, and allow different levels of access to data and metadata for different online services and platforms, typically with very little needed in terms of additional setup. Most importantly however, the service standards and protocols are open, understood, and widely accepted throughout the scholarly and academic library communities, ensuring a broad level of support and interoperability for PKP applications.

⁶ Dataverse Network. <http://thedata.org>

⁷ PLOS Article-Level Metrics: <http://article-level-metrics.plos.org/alm-info/>

⁸ Altmetrics: <http://www.altmetric.com/>; Orcid: <http://about.orcid.org/>

(b) **Enhancing Accessibility**

PKP promotes access to scholarly content in a number of ways. Interoperability via open standards and services, discussed above, is of key importance to accessibility: providing multiple access methods to content will of course increase exposure. In this fashion, journals may have their content harvested by OAI-capable metadata harvesters, can provide article DOI information to CrossRef, can deposit into PubMed's MEDLINE indexing service.⁹ All PKP applications are search-engine-friendly, and include special HTML "meta" tags that are used by Google Scholar to identify and present content properly. In addition, the application's HTML is written to specific standards, and special care is taken to ensure general accessibility across a wide variety of browsers and operating system.

(c) **Enhancing Openness**

Open Source Software

PKP software applications have always been released as open source software, under the General Public License.¹⁰

The software is free in two ways: It is free to download and use; and the source code is freely available to download, view, and modify. There are a number of reasons why PKP publishes these applications as open source software.

Firstly, our mandate to improve the access to and quality of scholarly research has been helped immensely by providing the software free of charge: Researchers from all over the world can download and use our software; in a very real sense, journals from Indonesia and Sri Lanka can operate on the same field (or quality of platform) as journals from the United States and Germany.

Secondly, our software has benefitted immeasurably from source code contributions from many, many members of the scholarly community. Almost all translations of the software have been contributed as code; bugs have been identified and in many cases fixed by community members; and new features (many in plugin format) have been contributed as well. Simply put, we would not have been able to attain the quality and breadth of our software without following an open source software community model.

Thirdly, while Open Access to scholarly research and open source software models are not necessarily explicitly interlinked, they do share some of the same philosophical underpinnings: Open Access to material; community collaboration; etc. Following an open Source licensing model makes as much sense as promoting an Open Access approach to scholarly research and to Open Science, more generally (cf. Willinsky 2005).

Open Community

PKP is a community project in many ways. The Project provides direct developer access to anyone, via the PKP support forums and wiki. Anyone can

⁹ CrossRef: <http://www.crossref.org/>

¹⁰ Specifically, the GPL V2. See GNU Operating System.

register on the forum and interact directly with the PKP development team for technical support or development inquiries. The support forum also exists as an active venue for questions and conversations from editors, with questions ranging from accessibility, indexing and DOIs to how typical review and editorial workflows typically work.

This community engagement is international. PKP seeks to cooperate with different community partners around the world (including translators, developers, and supportive institutions in Europe, Africa, Latin America, China, and elsewhere). PKP has worked with INASP and other organizations in delivering scholarly publishing workshops in many parts of the world. It was worked with INASP to build portals such as African Journals Online and Asia Journals Online, which have created an online presence in Google Scholar and elsewhere for hundreds of titles.¹¹ In addition, many independent user and training groups have popped up throughout the world, operating without direct PKP support—for example, one partner in Spain has developed an entire OJS training program and support forum, while another, with IBICT in Brasilia has been offering workshops across the country for years.¹² That these community initiatives are blossoming internationally, and largely without direct support from PKP, is a welcome marker of success, and an indication of the broad acceptance of PKP software as a scholarly publishing standard internationally.

Open Access

A key goal of PKP is to promote Open Access to scholarly research. As such it is part of growing Open Access movement. Open Access was initially, at the turn of the century, a radical challenge to the old print model, but it is now increasingly embraced not just by small independent scholar-publishers, where it got its start, but by the largest of scholarly publishing corporations, just as it is being supported by government legislation requiring Open Access for funded research, with every sign that Open Access may well become the norm for publishing research (cf. Laakso and Björk 2012; Björk and Peatau 2012). With the development of mega-journals, such as *PLoS One* publishing tens of thousands of Open Access articles a year, and increasing use of “article processing fees” to guarantee that Open Access can be a source of revenue, the momentum and incentive is transforming the industry (cf. Frank 2012). While PKP continues to largely serve smaller Open Access journals operated by scholar-publishers, efforts are underway to adapt its approach to make the mega-journal model among the options that it offers to the academic community.

Open Education

One of the challenges for sustaining, enhancing, and increasing Open Access is lack of professional publishing experience among many in the growing community. A new initiative of PKP is the development of tuition-free, open, online training courses in the use of PKP software and online publishing and management

¹¹ African Journals Online: <http://www.ajol.info/>; Asia Journals Online: <http://www.asiajol.info/>

¹² OJS.es: <http://www.ojs.es/>; IBICT: <http://www.ibict.br/>

skills.¹³ Working in conjunction with partners such as the International Network for the Availability of Scientific Publications (INASP) and the Publishing Studies Department at Kwame Nkrumah University Of Science and Technology (KNUST) in Ghana, this new education program will help build local capacity for online publishing and sustainable Open Access.¹⁴

(d) **Enhancing Knowledge**

The fundamental purpose of the Public Knowledge Project is to enhance the quality of scholarly communication and global knowledge sharing. By providing free and open source tools for professional monograph, journal, and conference management, PKP has enabled scholars from around the world to launch their own online publications with little or no money, and to build scholarly communities around their areas of interest, and to share the results of their research with all (cf. Willinsky and Mendis 2007).

Discussions of the developing world and Open Access often revolve around making research from the leading publications from the developed world more widely available. Programs such as HINARI, AGORA, and OARE have made significant gains in this area.¹⁵ While this is no doubt important, it is equally important for researchers in the developed world (and elsewhere) to hear the voices from the South. In the past, a leading publisher may have rejected the knowledge generated by a Ghanaian researcher because the significance of her work was not understood or valued. She could instead publish it in a local print-based research publication, but it would have a very limited readership, with physical copies not making it far beyond her country's borders. With the increasing availability of the Internet in Africa, although still a challenge, and the existence of free and open source tools for professional publishing, she has new options. Locally produced online journals, with a global community of editors, authors, and reviewers are increasingly available as a forum for her work, and where a suitable local option doesn't exist, she can now choose to collaborate with colleagues to start her own online journal.

Conclusion

The Public Knowledge Project has worked hard with a good number of organizations and institutions, editors and publishers, over the course of the last decade-and-a-half to increase the options and alternatives available to the global community of scholars and researchers. In the face of this new publishing medium that has transformed so many aspects of communication, and with even more

¹³ PKP School: <http://pkpschool.org>

¹⁴ INASP: <http://www.inasp.info/>; KNUST: <http://www.knust.edu.gh/pages/>

¹⁵ Research4Life: <http://www.research4life.org/>

changes clearly in the offing, it is too early to know or even predict what models and methods are going to prevail as the digital era of scholarly communication continues to unfold. Our project has always been to demonstrate ways in which these new directions and opportunities might uphold long-standing historical principles of openness, community, cooperation, experimentation, and questioning that continue to underwrite the work of research and learning. The continuing success of this work relies not only on the open nature of the project, but on the passion and interests of this larger community in their desire to contribute ideas and knowledge, as well as the always appreciated instances of well-formed code.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven, USA: Yale University Press. Available at: http://cyber.law.harvard.edu/wealth_of_networks/.
- Björk, B.-C., & Paetau, P. (2012). Open Access to the scientific journal literature—status and challenges for the information systems community. *Bulletin of the American Society for Information Science and Technology*, 38(5), 39–44. doi:10.1002/bult.2012.1720380512.
- Edgar, B.D., & Willinsky, J. (2010). A survey of scholarly journals using Open Journal System. *Scholarly and Research Communication*, 1(2). Available at: <http://src-online.ca/index.php/src/article/view/24/41>.
- Frank, (2012). Megajournals. *Trading Knowledge*. Available at: <http://occamstypewriter.org/trading-knowledge/2012/07/09/megajournals/>.
- Laakso, M., & Björk, B.-C. (2012). Anatomy of Open Access publishing: a study of longitudinal development and internal structure. *BMC Medicine*, 10(1), 124. doi:10.1186/1741-7015-10-124.
- Suber, P. (Ed.), (2012). What is Open Access? *In Open Access*. Massachusetts (pp.1–27). USA: MIT Press.
- Willinsky, J. (1999). *Technologies of knowing: a proposal for the human sciences*. Boston, Mass.: Beacon Press.
- Willinsky, J. (2005). The unacknowledged convergence of open source, Open Access, and Open Science. *First Monday*, 10(8). Available at: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1265/1185>.
- Willinsky, J. (2009). The stratified economics of Open Access. *Economic Analysis and Policy*, 39(1), 53–70.
- Willinsky, J., & Mendis, R. (2007). Open Access on a zero budget: a case study of Postcolonial Text. *Information Research*, 12(3). Available at: http://www.aughty.org/pdf/postcolonial_text.pdf.