Building Institutional Repositories in Theological Libraries

by Clifford B. Anderson and Charlotte Lew, Vanderbilt University; Ed Warga, Texas A&M University, Corpus Christi

INTRODUCTION

Institutional repositories are mainstays of the open access movement. However, theological schools have been slower to adopt them than other institutions of higher education. We believe that theological schools may be hesitant to build institutional repositories because they do not fully understand the principles of open access, have not conceptualized workflows to populate repositories, and have not developed effective communities of practice. In this essay, we try to provide perspectives on these three sticking points.

A BRIEF HISTORY OF OPEN ACCESS

Let's begin by providing a definition of open access. Peter Suber, a doyen of the open access movement, has described open access literature as "digital, online, free of charge, and free of most copyright and licensing restrictions."[1] In similar fashion, the Scholarly Publishing and Academic Resources Coalition (SPARC) defines open access as "the free, immediate, online availability of research articles coupled with the rights to use these articles fully in the digital environment."[2]

There is ambiguity at the heart of the open access movement. Does "open" mean simply the freedom to read online or does "open" imply other rights? According to Peter Suber, a distinction may be drawn between "gratis" and "libre" open access.[3] The first allows users to read materials without cost online whereas the second permits them to rework, translate, redistribute, and query materials computationally, among other things. There is ongoing conversation among experts about whether only the "libre" form genuinely qualifies as open access.[4]

We are experiencing a period of major innovation in the business models of open access. Currently, there are two primary business models, termed "green" and "gold" open access. The "green" model presupposes the traditional publishing process, but depends on authors retaining sufficient rights to "self-archive" or post some version of their works in institutional or disciplinary repositories. The "gold" model assumes that authors subvent the cost of publishing by paying fees to make their articles available online. These charges are commonly denominated article processing charges (APCs). There are also so-called "hybrid" journals, which continue to depend on subscription income from libraries and other institutions, but which allow authors to selectively make articles openly available upon payment of an APC.

We should also note that open access advocacy belongs, broadly speaking, to the field of scholarly communications. The Association of College & Research Libraries (ACRL) defines scholarly communications as "the system through which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community, and preserved for future use."[5] To understand how we arrived where we are today, we will take a brief journey through time and revisit some highlights in the history of scholarly communications, as adapted from Peter Suber's Open Access Timeline.[6]

The contemporary form of scholarly publishing first developed in the mid-seventeenth century. The English journal *Philosophical Transactions of the Royal Society,* which began publishing in 1665, is commonly credited as being among the first scholarly journals of the modern era.[7] The system of reporting academic findings via academic journals persisted more or less unchanged until the late twentieth century. Tim Berners-Lee's invention of the World Wide Web in 1989 proved the decisive event that initiated a revolution in the paradigm of scholarly publishing.[8] The newfound ability to instantaneously disseminate research worldwide without incurring the high cost of print publication begged the question why academics would continue to communicate with one another in print through traditional scholarly journals.

Forward-thinking academics began exploring other means for disseminating research in the early 1990s. During this era, the concept of a preprint server developed.[9] In the days before the web, authors shared printed drafts of articles with one another to get feedback and to anticipate lines of criticisms. The advent of preprint servers like ArXiv (1991) and SSRN (1994) permitted authors to continue this tradition in digital form. Of course, preprint servers also made these drafts widely available to the public, raising thorny questions about copyright. Early advocates of open access models formulated the concept of self-archiving, carving out rights in publication agreements for authors to make draft versions of their works accessible online. The Association for Research Libraries (ARL) also created SPARC during the late 1990s to advocate for open access to research literature.

In the 2000s, the movement toward open access began to pick up steam and many of the tools and concepts we associate with the movement today came into existence. Among the key innovations during this period was the formulation of the Creative Commons (CC) licenses in 2002. The CC licenses allow ordinary creators without legal backgrounds to select and publish materials with permissive agreements, reducing or eliminating legal frictions that impede distribution and reuse. Government agencies also played a vital role during this period by mandating open access to research outputs and by starting large-scale federal initiatives like PubMed Central (2002). Key tools for securing and discovering open access materials became available in the 2000s, from SPARC's Author's Addendum (2005), which supplies boilerplate legal language for authors to use when signing publication agreements in order to retain their freedom to distribute materials openly, to Project RoMEO at the University of Nottingham, which provides a database of publishers' open access policies. Open source software such as DSpace, ePrints, and Fedora Commons also made it feasible for universities to set up institutional repositories without the involvement of commercial vendors. University faculties attempted to shift the preference toward open access publication by passing so-called "institutional mandates," which require members of faculty to publish versions of their articles in open access unless they seek and obtain waivers from the institution. The University of Southampton maintains a project called the Registry of Open Access Repository Mandates and Policies (ROARMAP) to track such mandates internationally.[10]

Traditional publishers did not stand idly by during this period of innovation in scholarly communications. Major corporate publishers began to offer some form of self-archiving option or hybrid open access model to authors. Of course, these publishers also tried to thwart forms of open access publication they perceived as injurious to their financial bottom lines.

At Vanderbilt University, we have no open access policy requiring faculty to self-archive their scholarly articles, so the library staff is designing a faculty works archiving service to entice faculty to self-archive in DiscoverArchive, our Institutional Repository. What we have dubbed the "Faculty Works Archiving Service" includes:

1. Professional bibliography work where subject librarians and staff create a bibliography of faculty works encoded as structured data with links to the version of record and open access versions;
2. Rights and permissions checking;
3. Persistent open access in our institutional repository with links to the version of record.

This service, which is still in development, began in conjunction with the law library on behalf of law faculty. Library staff created a mediated deposit workflow to support this initiative. All faculty need to do is give the library permission to archive works on their behalf, and librarians do the rest. This workflow is now being implemented and improved upon in the divinity library by the divinity library staff, as discussed in the next section.

AN OPEN ACCESS WORKFLOW

The Vanderbilt Divinity Library staff initiated a faculty publications archiving project to build an institutional repository in April 2014. In the first phase of the project, the institutional repository coordinator provided Zotero training to five divinity library staff to create a comprehensive bibliography of divinity faculty publications. The core faculty from the Master of Divinity and Master of Theological Studies programs, a total of 35 members, was included in this project. Working through the permission process and maintaining a database of results took approximately ten to fifteen hours a week of the designated staff member.

The flowchart in Figure 1 (next page) is not yet set in stone as the process continues to unfold. The effort of utilizing library staff as mediators to build a collection of faculty works in the institutional repository for a theological institution is challenging. The IR team meets regularly for problem solving and strategic planning. Seeing the evolving workflow, the team expects more challenges to manifest themselves over time.

The divinity school's website where the faculty post and update their CVs is a great resource to rely on in terms of collecting faculty publications. Despite a consensus to focus the effort on getting journal article permissions, library staff also took time to create a comprehensive list of faculty publications. As a result, books, book chapters, book reviews, dictionary and encyclopedia articles, newspaper articles, blog postings,

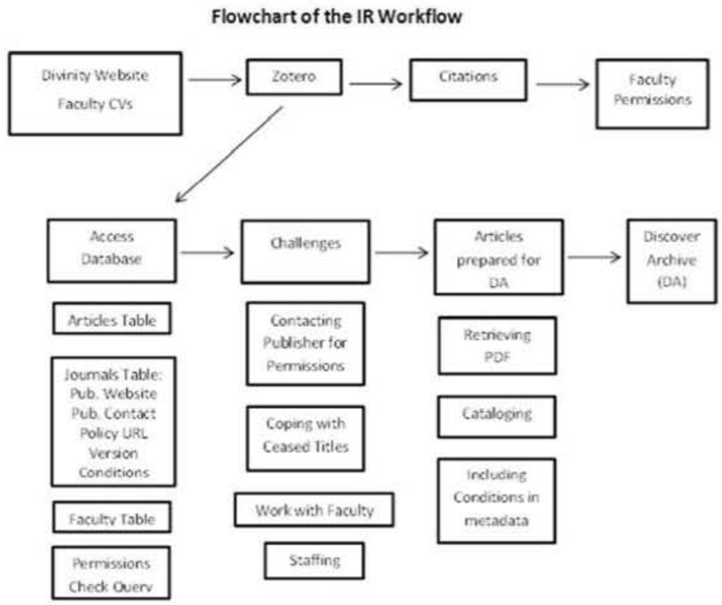

Figure 1: Flowchart of the IR Workflow

and webpages are all contained in the bibliography alongside journal articles. After building the bibliographies in Zotero, we informed the faculty about the IR program and endeavored to bring them on board. With strong support from the Dean of the Divinity School, nineteen permissions were received, which represents over 50% of the faculty.

When collecting publications in Zotero, surprisingly numerous errors were found on the CVs. The errors range from incomplete citations, inaccurate pagination, faulty publication dates, or, worse, incorrect titles. Being meticulous with the list of publications helps to incentivize faculty participation. Several faculty members expressed gratitude that their CVs had been corrected and omissions filled in.

What's the best way to work with faculty? Work meticulously on citations. Keep yourself available for questions. Wait on the harvest with patience and perseverance. Faculty can show their enthusiasm about the project and promise to deliver the permission when they are informed about the project; nevertheless, they may need several reminders before permissions are granted.

Microsoft Access is a great tool to keep track of the data retrieved from the publishers. Three tables are incorporated in the database: article, faculty, and journal. When needed, queries can be generated to filter or summarize data from one or multiple tables. The information whether the article has had its copyright cleared, has been acquired, and is ready for deposit are all available in the article table. The permission and authorized faculty name can be found in the faculty table. Most of the data is in the journal table, which records ISSN, publisher's website and contact information, and copyright policy URL and conditions of the policy. Currently, the database contains more than 850 articles in 320 journal titles which require additional research about open access and self-archiving policies.

There are two beneficial resources to use when researching for policies. Ulrichsweb, a global serials directory, brings together all the latest bibliographic and publisher information in one location. It is so efficient that details of the publisher's website, contact information, publication status, and title history are all at your fingertips. However, some downsides of Ulrichsweb exist. The updating of data may not happen in a timely matter. It is disappointing when the link to the publisher's website is obsolete or broken. Regardless, it does save time and frustration when searching. Another resource is SHERPA/RoMEO, which displays the copyright and open access self-archiving policies of academic journals. It uses a color-coding scheme to indicate what self-archiving policies the publisher offers to the authors: white for ungraded titles, yellow for titles that allow the archiving of preprints, blue for those permitting post-prints, and green for titles that permit published versions to be archived online. This is another time-saving tool that tremendously reduces the workload in digging out the policies from the publisher's website. Unfortunately, SHERPA/RoMEO may not always be up-to-date with current policies. Several titles listed as not supported by SHERPA/RoMEO were spotted to have archiving policies on the publishers' websites. If there is disagreement between self-archiving policies stated on publishers' websites and on SHERPA/RoMEO, we also need to clarify that discrepancy.

The development of this evolving project heavily hinges on faculty participation and publishers' assistance. The process cannot move forward without responses from faculty and publishers. Moreover, we still have to consult with Vanderbilt's Office of General Counsel about how to handle the question of orphan works for which no current publisher is available to handle our permission requests. It goes without saying that staffing is another issue since the process can be tedious and intensive. More staff time will be necessary to cope with the demands of sustaining the project over time.

The terms publishers use to describe the different publication versions are inconsistent. Here is an example of terms different publishers use to describe the post-print version: Taylor and Francis uses author's accepted manuscript, but Emerald allows the final version of articles. It takes extra effort to clarify the terms with the publishers. Emails, forms on the websites, and phone calls can be utilized to contact publishers for the self-archiving policies. However, not all the publishers provide contact information on their website (assuming a website even exists). One hundred thirty publishers have been contacted and forty-four have responded to date. The majority of those who responded granted the permission during the same week they received the request. Our inquiry emails are sometimes relayed to different people, from administration assistants to editors or even to legal officers. The request for Vanderbilt Divinity blog *Reading This Book* was referred to six people before permission was delivered two months after we first sent an email. For publishers from whom we have not received permission, we envision two additional tasks: locating contact information from the printed journals where websites do not exist and resending our requests to the publishers who have not responded to our initial inquiries. Before faculty publications may be deposited into the IR, a PDF of the published version has to be retrieved; in many cases, this means requesting and digitizing print copies from faculty. Furthermore, articles have to be cataloged with rights metadata indicating the permissions agreements.

DiscoverArchive serves as a showcase for the work of the Vanderbilt community. All content is indexed by Google and Google Scholar. Patrons can easily access the content from a Google search. DiscoverArchive not only provides free access to the work but also preserves the content into the future. A total of 216 divinity faculty papers are available from DiscoverArchive after two years of work.

This project serves as a catalyst for conversation in order to achieve the goal to build alliances and to promote open access publications. Educating faculty on the benefits of open access is crucial to the success of the project. Hopefully this established model can be extrapolated to other libraries and institutions.

A COMMUNITY OF PRACTICE

In this final section, we reflect on the possibility of developing a community of practice within ATLA member institutions to support open access and scholarly communications.

Let's start by providing a working definition of a "community of practice." According to Etienne C. Wenger and William M. Snyder, "In brief, they're groups of people informally bound together by shared expertise and passion for a joint enterprise...."[11] Unlike traditional committees, communities of practice are self-organizing and fluctuate organically in membership. Managers do not typically assign members to these groups, though they may support their activities in other ways. Finally, communities of practice may also draw together subject matter specialists from different functional areas. As Wenger and Snyder observe about a successful community of practice, "The participants in these communities of practice were learning together by focusing on problems that we directly related to their work."[12]

In many larger academic institutions, a community of practice has developed to support open access. Members of this community may come from across the university and include librarians along with faculty. There may also be administrative professionals in the Office of General Counsel or the technology transfer office with interest in open access publication and open source business models. While this community may not meet regularly together, a good scholarly communications office will help to foster connections between them and assist members of the network with learning the trends in open access publication models.

A challenge for many smaller member institutions is that they frequently do not have connections with university libraries. That means that they do not usually have direct access to a scholarly communications officers to advise and help to implement an open access program. As we saw in the second section of this paper, the open access initiative at the Vanderbilt Divinity School Library brings together librarians across the university library system. In addition, these initiatives rely in part on experience and feedback from librarians, staff, and members of faculty in other divisions of the university.

Without an institutional community of practice, librarians at smaller institutions must master the arcane rules and practices of scholarly communications by reading the professional literature and by sharing ideas at conferences, webinars, and social media. This mode of professional development can prove difficult, especially in light of all the other duties that staff members at these libraries must fulfill. In "The Future of the Small Theological Library," Myka Kennedy Stephens eloquently describes the challenges of keeping pace with professional development. As she argues, it is not simply sufficient *to translate* trends from larger university libraries; it's also necessary *to understand* and to move beyond them: "...We must start training ourselves to be more forward thinking, mindful of current trends yet looking beyond them at the same time."[13] Certainly, this is the right perspective for fostering scholarly communications in the ATLA. ATLA member institutions cannot simply imitate trends at large research universities; they need to adapt them and transform them to serve the goals of their particular institutions. A community of practice in the ATLA could foster that process.

There already exists something of a community of practice around open access in the ATLA. David Stewart, Director of Libraries at Bethel University, and colleagues launched *Theological Librarianship* as an open access publication in 2007. Gary F. Daught, Director of Libraries at Milligan College (TN), actively blogged about open access in theology and religious studies from 2011 to 2015.[14] There are frequent conversations on the ATLANTIS listserv about open access; the conversation also benefits from the participation of Kevin L. Smith, former Director of Copyright and Scholarly Communication at Duke University, who regularly weighs in on tricky questions of copyright. Other grassroots initiatives to promote open access, including multiple institutional repositories, exist across ATLA member institutions. These efforts have also produced an official movement toward fostering scholarly communications within the ATLA. In 2015, the ATLA Board of Directors adopted a strategic plan stating that its "core purpose" is "To promote worldwide scholarly communication in religion and theology by advancing the work of libraries and related information providers."[15]

What might the ATLA do to promote scholarly communication among its member institutions and more broadly? Here are a few ideas, roughly in order of complexity and challenge.

First, the ATLA might produce a model open access agreement for publishers who do not presently offer open access publication options. While large corporate publishers like Elsevier, Sage, and Wiley offer well-defined (if not necessarily generous) open access options, publishers such as scholarly societies and theological schools with one or two journals often do not articulate open access policies at all. Our intuition is that these smaller publishers are not against open access; they simply need help to understand the movement. Proffering a model agreement with annotations to indicate the implications of offering certain open access options might encourage smaller, less experienced publishers to experiment with open access business plans.

Second, the ATLA might assist smaller publishers with publicizing their open access options. As mentioned previously, the best source of information for publishers' policies is the SHERPA/RoMEO site, maintained by the University of Nottingham in the United Kingdom.[16] While journals like *Religious Studies Review* and *Theology Today* have their self-archiving policies recorded in SHERPA/RoMEO, many others do not. If a publisher is not listed on the site or is listed as "ungraded" (i.e., with no reported policy), scholarly communications librarians will need to research their policies, generally by writing their editors for clarification. This process of soliciting information greatly slows down the clearance process. A useful project would be to compare the list of journals in ATLAS with the SHERPA/RoMEO registry and to encourage publishers of "ungraded" journals to submit their policies. Of course, this would ideally take place after those journals adopted model open access policies.

Third, the ATLA could provide education and recommendation about preferred open access licenses. A wide spectrum of options exists for making materials "open access," ranging from providing online access to articles without toll gates (but also without licenses) to dedicating materials to the public domain. The Creative Commons licenses help to fill the spectrum by providing very granular options.[17] While all Creative Commons licenses require licensors to attribute the original creators (and other downstream modifiers, if any), it is also possible to pair attribution (BY) with additional restrictions such as ShareAlike (SA), NonCommercial (NC), and NoDerivatives (ND). How should one select among the combinatorial possibilities? If the ATLA recommended a particular license for open access articles, say CC-BY, that could inform policy at ATLA member institutions.

Fourth, the ATLA might lead an initiative encouraging faculty members, students, librarians, and other alt-ac staff members to sign up for ORCID. ORCID is a nonprofit that provides scholars with a unique identification number that they can use to associate themselves with their research output. ORCID is similar to other name authority schemes like the Library of Congress Authorities or the Virtual International Authority File (VIAF) with the difference that it cedes control to researchers to describe themselves. By entering researchers' ORCID when adding materials into the institutional repository, institutions make it easier to identify, discriminate, and track individuals' scholarly production.[18]

Finally, the ATLA could connect theological institutions that have existing institutional repositories with national scholarly communications initiatives like SHARE. SHARE emerged in response to the Office of Science and Technology Policy memorandum of February 22, 2013, mandating that federal agencies with budgets over $100 million develop public access policies for their research outputs.[19] A goal of SHARE is to link institutional repositories together by providing common metadata schemas. If lots of theological institutions participated, we would be able to gain perspective on our research activities as a whole. Right now, the majority of religious studies and theology entries in SHARE appear to come from larger journal publishers and from data sources like CrossRef and DataCite. By encouraging schools to connect their institutional repositories with SHARE, we would likely develop a more balanced picture of theological research events at ATLA member institutions.

There are, of course, many other ideas that could foster the institutional repository movement in the ATLA. These proposals are meant only to provide a starting point for developing a community of practice around open access. None of them carry price tags apart from the (significant) personnel costs of coordinating and carrying them out. From our perspective, the primary challenge to developing open access initiatives at theological schools is less financial than cultural or social. Returning to our theme above, many theological librarians do not have colleagues to assist them with setting up and articulating the goal and purpose of scholarly communications initiatives. Developing a community of practice around open access, institutional repositories, and scholarly communications would encourage and empower more theological librarians to become involved in these contentious, but critically important, issues for our discipline.

WORKS CITED

Association of College & Research Libraries (ACRL). "Principles and Strategies for the Reform of Scholarly Communication 1." Text, 2006. http://www.ala.org/acrl/publications/whitepapers/principlesstrategies.

ATLA Board of Directors. "ATLA Strategic Plan," 2015. https://www.atla.com/about/who/Pages/ATLA-Strategic-Plan.aspx.

Daught, Gary, and Peter Webster. "Omega Alpha." *Omega Alpha Open Access*. Accessed September 9, 2016. https://oaopenaccess.wordpress.com/.

"Licensing Types." *Creative Commons*, 2016. https://creativecommons.org/share-your-work/licensing-types-examples/.

"ORCID," 2016. http://orcid.org/.

"Philosophical Transactions of the Royal Society of London," 2016. http://rstl.royalsocietypublishing.org/.

Poynder, Richard. "Open and Shut?: The Open Access Interviews: Paul Royster, Coordinator of Scholarly Communications, University of Nebraska-Lincoln." Open and Shut?, 2014. http://poynder.blogspot.com/2014/08/the-open-access-interviews-paul-royster.html.

"Preprint." *Wikipedia, the Free Encyclopedia*, September 2016. https://en.wikipedia.org/w/index.php?title=Preprint&oldid=738269465.

"ROARMAP," 2016. https://roarmap.eprints.org/.

"SHARE." SHARE, 2016. http://www.share-research.org/.

SPARC. "Open Access." Accessed September 8, 2016. http://sparcopen.org/open-access/.

Stephens, Myka Kennedy. "The Future of the Small Theological Library." *Theological Librarianship: An Online Journal of the American Theological Library Association* 9, no. 1 (2016): 28–32. https://theolib.atla.com/theolib/article/view/411.

Suber, Peter. "Gratis and Libre Open Access," 2008. https://dash.harvard.edu/bitstream/handle/1/4322580/suber_oagratis.html.

———. "Open-Access Timeline (Formerly: FOS Timeline)," 2009. http://legacy.earlham.edu//~peters/fos/timeline.htm.

———. "Very Brief Introduction to Open Access," 2004. http://legacy.earlham.edu/~peters/fos/brief.htm.

Wenger, Etienne C., and William M. Snyder. "Communities of Practice: The Organizational Frontier." *Harvard Business Review*, 2000. https://hbr.org/2000/01/communities-of-practice-the-organizational-frontier.

World Wide Web Foundation. "History of the Web," 2015. http://webfoundation.org/about/vision/history-of-the-web/.

ENDNOTES

[1] Peter Suber, "Very Brief Introduction to Open Access," 2004, http://legacy.earlham.edu/~peters/fos/brief.htm.

[2] SPARC, "Open Access," accessed September 8, 2016, http://sparcopen.org/open-access/.

[3] Peter Suber, "Gratis and Libre Open Access," 2008, https://dash.harvard.edu/bitstream/handle/1/4322580/suber_oagratis.html.

[4] Richard Poynder, "Open and Shut?: The Open Access Interviews: Paul Royster, Coordinator of Scholarly Communications, University of Nebraska-Lincoln," *Open and Shut?*, 2014, http://poynder.blogspot.com/2014/08/the-open-access-interviews-paul-royster.html.

[5] Association of College & Research Libraries (ACRL), "Principles and Strategies for the Reform of Scholarly Communication 1," Text, (2006), http://www.ala.org/acrl/publications/whitepapers/principlesstrategies.

[6] Peter Suber, "Open-Access Timeline (Formerly: FOS Timeline)," 2009, http://legacy.earlham.edu//~peters/fos/timeline.htm.

[7] "Philosophical Transactions of the Royal Society of London," 2016, http://rstl.royalsocietypublishing.org/.

[8] World Wide Web Foundation, "History of the Web," 2015, http://webfoundation.org/about/vision/history-of-the-web/.

[9] "Preprint," Wikipedia, the Free Encyclopedia, September 2016, https://en.wikipedia.org/w/index.php?title=Preprint&oldid=738269465.

[10] "ROARMAP," 2016, https://roarmap.eprints.org/.

[11] Etienne C. Wenger and William M. Snyder, "Communities of Practice: The Organizational Frontier," *Harvard Business Review*, 2000, 139, https://hbr.org/2000/01/communities-of-practice-the-organizational-frontier.

[12] Ibid., 143.

[13] Myka Kennedy Stephens, "The Future of the Small Theological Library," *Theological Librarianship: An Online Journal of the American Theological Library Association* 9, no. 1 (2016): 30, https://theolib.atla.com/theolib/article/view/411.

[14] Gary Daught and Peter Webster, "Omega Alpha," *Omega Alpha Open Access*, accessed September 9, 2016, https://oaopenaccess.wordpress.com/.

[15] ATLA Board of Directors, "ATLA Strategic Plan," 2015, https://www.atla.com/about/who/Pages/ATLA-Strategic-Plan.aspx.

[16] SHERPA stands for "Securing a Hybrid Environment for Research Preservation and Access" and RoMEO stands for "Rights MEtadata for Open archiving."

[17] "Licensing Types," Creative Commons, 2016, https://creativecommons.org/share-your-work/licensing-types-examples/.
[18] "ORCID," 2016, http://orcid.org/.
[19] "SHARE," SHARE, 2016, http://www.share-research.org/.