

Competing Information Realities: Digital Libraries, Repositories, and the Commons

Presenters: Donald Kraft, Louisiana State University & Editor, *JASIST*; Edie Rasmussen, University of British Columbia, Samantha Hastings, University of South Carolina & Editor, *ASIS&T Monograph Series*; and Anita Coleman, University of Arizona and Editor, *dLIST*.

Organizer: Anita Coleman (asc@u.arizona.edu)

Sponsor: SIG DL

Introduction and Scope

Scholarly communication today faces unprecedented social and economic challenges coupled with the unbridled promises of new technologies that purport to solve many of its problems. Most of the problems are largely due to the serials price crisis in science, technology, and medicine, but others have also existed for awhile now such as tedious publication delays and criticism of the anonymous peer review process (Bachrach, et al, 1998; Mannheim, 2000). Facilitated by advances in information and communication technologies digital libraries at first (Schatz, et al, 1994; Atkinson, 1996; Harter, 1997), and more recently digital repositories (Atkinson, 2003; Barton, 2003) have emerged as alternative scholarly communication and publishing media. Heery and Anderson (2005) provide a review of digital repositories but very simply, digital repositories are unlike digital libraries such as Perseus, in that repositories provide a submission mechanism whereby scholars can deposit an electronic copy of their work at the time or soon after creation. Digital repositories are also called open access archives because of the lack of tolls, fees, or other legal and economic restrictions to access the content they make available (Ginsparg, 1996). For Lynch (2003), repositories are an essential infrastructure for scholarship in the digital age but for Tennant (2005) they merely collect “grey literature.” This situation is further exacerbated by references to the “commons” whose metaphorical meanings (like libraries and repositories) also range from the public domain to ideas far larger than that. Using the open source software model of knowledge production as a main example, Benkler (2002) argues for understanding modern information and knowledge production as a commons-based peer production model; this is also a view that makes sense when one considers the popularity of wikis or the global open access archives that have been operating in disciplines like high energy physics and computer science for more than a decade now. Even more interestingly, the Conservation Commons (2006) is using the idea of the commons to solve problems of biodiversity and the Science Commons (2006) is promoting the easing of barriers for scientific information of all types. The commons has enormous implications for research and teaching in the communities that make up the information sciences. Yet, the overwhelming majority of the ISI-ranked IS/LS journals today remain closed, there appears to be little innovation in our disciplinary scholarly communication system, there is only modest research from the information sciences community about the 'commons' and its potential for innovating information sciences research, including impact on digital libraries and digital repositories remains unexplored.

The goal of the panel is to explore the concept of the commons by framing it in the context of scholarly communication while also honing our understandings about digital libraries and repositories as technologies and socio-cultural artifacts. Panel members will uncover the pros and cons of the commons for LIS research and scholarly communication by describing the cognate and competing extant information realities. Edie Rasmussen will discuss the role of digital libraries in the commons. Anita Coleman, dLIST editor, the first open access archive for the information sciences will present her latest research about open access archives and the commons. Donald Kraft, Editor-in-chief of *JASIST*, will share his experiences editing a peer-reviewed ISI-ranked journal. Samantha Hastings, editor of *ASIS&T monographs* will share book publishing plans and concerns.

Digital Libraries and the Commons (E. Rasmussen)

The digital library, with its emphasis on selection, organization, community, and networking, has an obvious role to play in the dissemination of the Commons. Krowne (2003) talks of the motivation and role of “Commons-based peer production” in the creation of digital library content. In other cases, the content of digital libraries of the Commons is being created institutionally. In the fall of 2004 Google announced the Google Print program, later called Google Book Search, which comprises a complex Google vision and alliance with publishers and with libraries to make books discoverable by the general public searching/googling the web. This initiative involves the University of Michigan, Harvard, Stanford, Oxford, and the New York Public Library, and is allowing the scanning of all books held including those within copyright. The Google vision has unleashed criticism, acclaim, and even competition. The Open Content Alliance (Tennant, 2005) was subsequently announced by The Internet Archive, Yahoo!, the California Digital Library, and others, with a focus on out-of-print materials to avoid the copyright challenges that Google is facing. In a talk titled “Triangulation 1: Google Print”¹, Lawrence Lessig, John C. Dvorak, and Leo Laporte highlight a number of important issues. They estimate that there are 32 million books in the US libraries, of which 5 million are in the public domain, and 3 million still in print, leaving 24 million in copyright and not in print. This results in a large volume of material that has no continuing commercial value, and yet which is not freely available. They note that changes in our approach to copyright are needed to address these problems. But enabling access through digital libraries is just one step. There are also many issues around the use of the materials. What do we know about digital access to millions of books? How can they be used, how can the user find the right texts and the right information within those texts? This is a new environment for digital access with many unanswered questions. These are some of the issues that I will discuss as I focus on sketching the role of digital libraries in the Commons based on current practices and research..

Questions to the Audience:

Is the digital library the right model for managing the Commons? Does it correspond to the users’ perception of a library? Does it send the right message?

Often the problem is not material which is clearly in or out of copyright, but the vast quantity of material for which the copyright status is unknown or unclear. How can a digital library help to promote the use of this material in digital form, while offering rights management to the copyright holder?

Can we deliver massive amounts of full-text digital content in a meaningful, human-centered way? Have we adapted the technology to the content to make it not just available, but useable?

Brief speaker bio:

Edie Rasmussen is Professor and Director of the School of Library, Archival and Information Studies at the University of British Columbia. Prior to joining UBC, she spent 15 years on the faculty of the School of Information Sciences, University of Pittsburgh. She has also held faculty and visiting positions at Dalhousie University, Nova Scotia; Institut Teknologi MARA, Malaysia; Nanyang Technological University, Singapore, and Victoria University of Wellington, New Zealand, and professional positions with Agriculture Canada and Atomic Energy of Canada. Her research interests are in digital libraries and information retrieval of text and images.

¹ Available for download at <http://thisweekintech.com/tri1>

Digital Repositories and The Commons (A. Coleman)

Institutional repositories, open access archives, and data repositories are only a few types of digital repositories to which users can contribute content. Scholarly behavior, however, varies by disciplines and many studies continue to bemoan the slow growth of institutional repositories; apparently faculty in most disciplines, do not see themselves as active participants in repository use, much less it's design, development or growth. When they do use repositories to contribute content, faculty report facing formidable and time-consuming barriers. For example, a faculty who wishes to self-archive soon finds out that journal copyright transfer agreements and the intellectual property rights of authors are ambiguous. Scholars also appear to have more loyalty to the discipline rather than to their institution and hence are reluctant to participate in institutional repositories. Finally, some researchers have argued that institutional repositories, where each institution runs it's own archive are not economically sustainable. Many small institutions cannot afford even the minimal costs associated with establishing a repository. These and other problems can be sorted into four categories of concern: intellectual property rights, value-added services including economies of scale, sustainability, and scholarly information needs and behaviors. In Library and Information Science (LIS) a solution to many of these problems has been proposed - the development of an LIS commons through a society-led global scholarly communication consortium (Coleman and Roback, 2005). This panelist will present a conceptual overview of a commons for the information sciences. Results from a study of the self-archiving behaviors of LIS scholars (announced at ASIS&T 05 conference and conducted in November 2005), ongoing developments about the commons, and the related commons-based peer production of knowledge phenomenon, are used to define our disciplinary commons and the potential for innovating its scholarly communication system and research.

Questions to the Audience:

- 1) Self-archiving in OAI-compliant repositories appears to be a strategy for building the commons. Where is self-archiving and OAI covered in the curriculum? [or where should they be]
- 2) What is the self-archiving potential and promises for our discipline? [that is, how many self-archive, how often, what types of materials, what advantages does self-archiving in oai-compliant repositories confer?]
- 3) What are the advantages/benefits of disciplinary open access archives and aggregation services?
- 4) Do the information sciences need disciplinary open access archive and aggregation services?
- 5) Open source software is a commons-based peer production of knowledge model. What is a comparable commons-based peer production of knowledge activity in our communities' research, education and practice [hint: bibliometrics datasets, information retrieval research, and creation of information literacy tutorials would qualify to me – but what do you think?]

Brief speaker bio:

Anita Coleman is Asst. Professor in the School of Information Resources & Library Science, University of Arizona, Tucson which she joined in 2001. Before coming to Arizona, she was with the Alexandria Digital Library and prior to that in various technical positions in academic libraries. Her areas of interest are Knowledge Organization, Human Information Behaviors, and Scholarly Communication. She has been growing *dLIST* organically since 2002.

Scholarly Journal Publishing (D. Kraft)

Scholarly journal publishing has seen some major changes in the last few years and we shall look at a few key experiences as an editor in order to highlight them. Very simply, editors, along with a critical peer-refereeing system, review and certify the content submitted. Thus, editors and peer reviewers are important gatekeepers of information and knowledge disseminated in the relevant discipline(s).

Technological advancements, besides the movement towards the commons, for example, the open access of all peer-reviewed scholarly journal literature, are causing a stir in the way journals are edited. These include: 1) online journal article submission systems that enable editors and the editorial offices to accept, manage and track author submissions systematically, and 2) open access archives which enable authors to post their eprints at any time: at time of submission to the journal (this version is often called pre-prints to indicate the version has not yet been peer-reviewed) or after acceptance and publication. We shall focus on some of the concerns and issues these and other phenomenon such as the “publish or perish” syndrome raise for editors and scholarly journals in the rest of my talk.

Questions to the Audience:

1) Do journals matter? What is the value of a journal today? Michael Gorman has suggested that we don't need *ejournals*, they only serve editors, not readers who are better served by the disaggregation of journals. Readers are interested in articles only and should be able to put them together. But Carol Tenopir among others has elaborated (2006) on the value of journals. What do you think?

2) As communities with an ethos/culture of open access to scholarship, we would like to support not-for-profit scholarly society journal publishing and open access to the research outputs of our own disciplines. But clearly we cannot compete with commercial publishers; for example, we're good at content review and certification but we benefit from the fact that publishers have business planning expertise and product development skills that might not be as easy or feasible for us to duplicate. What are the strategies for acting on our values and being sustainable at the same time?

3) The so-called publish or perish syndrome has given rise to numerous new journals/publications. What are the pros and cons of the syndrome for authors, readers, and librarians?

4) How should the traditional, peer-reviewed journal change? Should it change?

5) To librarians: do you consider the immediate free availability of content in open access archives a good reason in itself to cancel a journal subscription? or this could be phrased as follows: if you knew a journal that actively encouraged and supported self-archiving in an open access archive, but the journal itself is a subscription journal, would you cancel your subscription?

Brief speaker bio:

Donald H. Kraft is currently a Distinguished Visiting Professor at the U.S. Air Force Academy, a professor in the Department of Computer Science, and an adjunct professor of library and information science at Louisiana State University (LSU). A former department chair in computer science at LSU and an ASIST past president his primary research interests lie in information retrieval models, and the application of artificial intelligence tools, such as fuzzy sets, rough sets, and genetic programming, to such retrieval models. He is the recipient of numerous awards: ASIST Best Information Science Book Award, ASIST Research Award, ASIST Watson Davis Award, LSU Distinguished Faculty award, LSU Community Volunteer Award. He is also an IEEE and an AAAS Fellow.

Scholarly Book Publishing (S. Hastings)

Our intellectual property and our intellectual, information commons, are at risk as more and more publishers become part of larger conglomerates (Munroe, 2006). Luckily for ASIS&T, our monograph series, is published by Information Today, Inc. (ITI). ITI still maintains an editorial staff and does not require camera-ready text. The largest challenge we face is how to move monographs into the digital environment and open source flexibility without losing the value-added perks of validation and reliability. Of course the economic argument for the portability of paper is a strong one but we must look ahead to what our future audiences might need from our scholarly content. I envision a different type of working environment where the scholarly and research endeavor is much more closely tied to cyber communities and personal digital libraries. If our users want to make annotations and share ideas in real time, then we need to think very carefully about our scholarly production, and how we can best build for future content, use and preservation.

The technological, economic, and legal challenges are large, but ultimately, our disciplinary norms and values are probably the critical factors that should and will drive our scholarly communication and publication systems. A proactive role includes exploring the nature and characteristics of a sustainable scholarly communication and publishing system for the information disciplines, engaging in a deeper understanding beyond mere productivity studies of our disciplinary scholars' information behaviors and needs, and boldly imagining a commons-based peer production model of scholarship as one strategy in the arsenal reshaping our cultural heritage institutions into critical information infrastructures (King, 2005).

Questions for the audience:

1. How many of you do the majority of your reading "on-screen"?
2. Does your institution include monographs in their digital repositories?
3. How do you communicate with other scholars in your field?
4. How many books do you purchase a year? How many scholarly books do you read a year?
5. What types of information do you seek in books?

Brief speaker bio:

Samantha K. Hastings is Professor and Director of the School of Library and Information Science at the University of South Carolina. She specializes in telecommunications, automated systems and networks, digital images, and evaluation research. She teaches Introduction to Digital Imaging, Advanced Topics in Digital Imaging, and Telecommunications for Information Professionals. She has various national and international publications and presentations in the field of Information Science. Dr. Hastings has an active role in American Society for Information Science and Technology and served as President of the Society in 2004. Her other professional affiliations are American Library Association for Library and Information Science Education, Museum Computer Network, and Texas Library Association.

Acknowledgments

Comments and questions by Tefko Saracevic (Rutgers) and Raym Crow (ARL SPARC) greatly helped in the development of this proposal. Thanks to them. Our thanks to the five anonymous reviewers who have encouraged us to focus and elaborate on the solution.

References

- Atkinson, Ross. (1996) Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone. *Library Quarterly* 66:239-65.
- Atkinson, R. C. (2003) A New World of Scholarly Communication. In *Chronicle of Higher Education*. November 07, 2003. <http://chronicle.com/free/v50/i11/11b01601.htm> (Accessed February 10, 2006).
- Bachrach, S., et al. (1998). Who owns scientific work? *Science* 281 (5382): 1451. Retrieved from EBSCOHOST (subscription database).
- Barton, M. R. & Walker, J. H. (2003) Building a Business Plan for DSpace, MIT Libraries' Digital Institutional Repository. In *Journal of Digital Information*. 4(2) Article No. 173, 2003-05-07. <http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Barton/barton-final.pdf> (Accessed February 10, 2006).
- Benkler, Y. (2002) Coase's Penguin, or, Linux and *The Nature of the Firm*. *Yale Law Journal* 112. Retrieved from Lexis Nexis Academic Universe (subscription database)
- Coleman, A. and Roback, J. (2005). Open access federation for Library and Information Science: DLIST and DL-Harvest. *D-Lib Magazine* 11 (12) December. <http://www.dlib.org/dlib/december05/coleman/12coleman.html> (Accessed February 10, 2006).
- Conservation Commons. (2006). <http://www.conservationcommons.org/>
- Creative Commons. (2006). <http://www.creativecommons.org/>
- DLIST. (2006). <http://dlist.sir.arizona.edu/>
- Ginsparg, P. (1996). Winners and losers in the global research village. In Conference held at UNESCO HQ, Paris, 19-23 Feb 1996, during session *Scientist's View of Electronic Publishing and Issues Raised*, Wed 21 Feb 1996. <http://people.ccmr.cornell.edu/~ginsparg/blurb/pg96unesco.html> (Accessed February 10, 2006).
- Harter, S. (1997). Scholarly communication and the digital library. *Journal of Digital Information* 1 (1). <http://jodi.tamu.edu/Articles/v01/i01/Harter/> (Accessed February 10, 2006).
- Heery, R. and Anderson, S. (2005). Digital Repositories Review. http://www.jisc.ac.uk/uploaded_documents/rep-review-final-20050220.pdf (Accessed February 10, 2006).
- King, J. (2005). Stepping Up: Shaping the Future of the Field. Keynote presentation at the 2005 ALISE Annual Conference, Boston, MA. <http://dlist.sir.arizona.edu/739/> (Accessed 13 February, 2006).

Krowne, Aaron (2003). Building a digital library the Commons-based peer production way. *D-Lib Magazine*, 9(10), <http://www.dlib.org/dlib/october03/krowne/10krowne.html>. ((Accessed February 13, 2006).

Lynch, Clifford. (2003). Institutional Repositories: Essential infrastructure for scholarship in the digital age. *ARL Bimonthly Report* 226: February. <http://www.arl.org/newsltr/226/ir.html>.(Accessed February 10, 2006).

Mannheim, A. (2000) When did peer review become anonymous? *Aslib Proceedings* 52 (8): 273-276.

Munroe, M. (2006). The Academic Publishing Industry: A Story of Merger and Acquisition. <http://www.niulib.niu.edu/publishers/> (Accessed 13 February, 2006)

Schatz, B., Bishop, A., Mischo, W., and Hardin, J. (1994). Digital library infrastructure for a university engineering community. In Digital Libraries 1994. *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, College Station, Texas, June 19-21, 1994. <http://www.csdl.tamu.edu/csdl/DL94/paper/schatz.html>. (Accessed February 10, 2006).

Science Commons. (2006). <http://www.sciencecommons.org/>. (Accessed February 10, 2006).

Tennant, R. (2002). Institutional Repositories. *Library Journal*, September 15. <http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA242297> (Accessed February 10, 2006).

Tennant, Roy (2005). The Open Content Alliance. *Library Journal*, <http://www.libraryjournal.com/article/CA6289918.html> (Accessed February 13, 2006).

Tenopir, C. (2006). The Value of the Container. *Library Journal*, <http://www.libraryjournal.com/article/CA6299865.html> (Accessed February 13, 2006)