Special Feature

Innovation Fair Abstracts

SPARC Open Access Meeting Kansas City | March 12 & 13, 2012

About the SPARC Meeting

SPARC (The Scholarly Publishing and Academic Resources Coalition) hosted its first North American meeting on Open Access in 2012. The first of its kind, the event expanded on the biennial SPARC Digital Repositories meeting. Convened by the organization since 2004 and hosted in the UK, Europe, Japan, and North America, these meetings have played an integral part in advancing the potential of open online repositories to expand the dissemination of scholarship and transform scholarly communication.

About the Innovation Fair

The Innovation Fair is a highlight of the biennial SPARC Open Access Meeting. Accepted participants provide rapid-fire presentations devoted to scholarly communications issues, sharing creative approaches to a diverse set of challenges. This year, a suite of librarians, technologists, research producers, and publishers presented, in no more than two minutes, innovative or creative approaches to: the use of open content, content discovery, value-added services, impact assessment, commercial innovation, and Open Access advocacy.

Abstracts © 2012 Abstract Authors. These abstracts are distributed under a Creative Commons Attribution 3.0 Unported License, which allows unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
001. Tapping Technical Services for Institutional Repository Engagement
Marianne Buehler, University of Nevada, Las Vegas

As library budgets shrink and journal articles assume the lion's share of purchasing, there are fewer books added to the collection. Subsequently, technical services have changing priorities that affect their work. The staff has a valuable skill-set that can benefit an institutional repository (IR), improving the efficiency of processing and archiving. As catalogers, they have metadata expertise and their descriptive and analytic abilities are useful for creating item records that optimize search engine retrieval. Other technical services personnel may check article copyrights, utilizing Sherpa Romeo or publisher websites, scan archival or other print materials, manipulate multimedia files, and upload a variety of intellectual content.

With numerous personnel involved and engaged in the IR, it is essential to have a workflow process that identifies what steps have been completed and by whom. At the University of Nevada, Las Vegas Libraries, the Technical Services staff and manager have successfully built and pilot-tested a Drupal software installation to manage communication and tracking for multiple projects. The IR manager alerts staff to new repository content, adding details in the item-type sections. Anyone who “touches” the project makes notes of copyright checking or other activities, such as adding metadata.

Communication is facilitated from the initial acquisition of content by the IR manager to the: Item copyright check, processing details, final deposit, and quality control, all in one location. Utilizing this primary communication tool, IR staff has embraced their value-added activities and roles in the open access enterprise.

Supplemental Content: View/Download Buehler Poster

002. Repositories for Social Science Data Collection
Leslie Carr, University of Southampton

Social scientists recognise the opportunity of the social Web (Twitter, Facebook, YouTube, email forums, etc.) as a source of data for research, but also the problem of the ephemeral nature of this data. We have been using a repository as mechanism for collecting such data for study over a long period, for preserving access to it, and for helping to analyse it. By embedding the repository into the research cycle from the very start, and by giving it a useful role in the research process, our aim is to encourage Open Access to the eventual research outcomes and the intermediate research data. The repository has been used by a research student who has collected 2.5 million tweets to study influence in the Twitter social network, and has deposited two OA preprints. It will be promoted to this year’s cohort of Masters students for their summer research projects as part of our institution’s new “Web Observatory” research agenda. This work complements our Research Data Repository activities, supported in the UK by JISC’s IDMB and DataPool projects.

An example of the problems that social scientists face in using here-today-gone-tomorrow web-based resources in their research, is described by the Data Curation Centre at Edinburgh University: http://www.dcc.ac.uk/news/challenges-managing-social-media-research-data-researchers-perspective.

Further details about our repository/social media work can be found at http://repositoryman.blogspot.com/2011/10/using-eprints-repositories-to-collect.html.

003. Springboard to Awareness: Student Publishing & Intellectual Entrepreneurship
Stephanie Davis-Kahl, Illinois Wesleyan University

Building on a challenging curriculum for students majoring in Economics, and featuring excellence in critical thinking and writing, the Undergraduate Economic Review (UER) immerses students in an online publishing experience that reframes and expands their understanding of research, peer review and audience beyond the classroom. At the heart of publishing the UER, a born-digital, open access student journal, is the notion of intellectual entrepreneurship— “educating individuals who utilize their intellectual assets to add to disciplinary knowledge and social good.” (Cherwitz)

A strong collaboration between the journal’s faculty advisor, the Scholarly Communications Librarian and the student editor-in-chief provides the human infrastructure for the journal, and each partner brings complementary knowledge and expertise to the table to facilitate a successful experiential project for students. With the successful launch of this online journal, several challenges abound: the rising number of submissions, finding adequate incentives to increase the pool of student reviewers, developing editorial talent over time, assessment of student editors’ and reviewers’ experiences, and marketing the journal effectively. Amid these challenges lies the educative imperative to keep Open Access tenets at the forefront of discussions and teachable moments with established and emerging scholars alike to prime them for further action and advocacy. Assessment of students’ understanding of Open Access is underway and will be reported on during this presentation.

Supplemental Content: View/Download Davis-Kahl Poster
Mel DeSart, University of Washington

What do you get when you combine:

- Tens of thousands of technical reports from academic library collections
- Engineering and government document librarians from over a dozen different institutions
- Two large academic library consortia
- A handful of different content digitizers
- Google
- Hathi Trust
- A cataloging/metadata creation operation
- Roughly 50 (and increasing monthly) entities of various kinds, both public and private, U.S. and beyond
- A goal and a grassroots effort to make openly accessible some of the most underutilized collections in academic libraries

You get TRAIL, the Technical Report Archive and Image Library: a collaboratively developed, openly accessible resource being used by people all over the globe. TRAIL began in 2006 as an initiative of the Greater Western Library Alliance and is now a Global Resource Network at the Center for Research Libraries. To date TRAIL has digitized and created metadata for over 28,000 U.S. government agency technical reports (well over 2.25 million pages), with more being processed daily. All are freely searchable and accessible via http://www.technicalreports.org.

For more information on TRAIL, see http://www.crl.edu/grn/trail.

TRAIL might simply be labeled as yet another digitization project, and at its core that's what it is. But that's not WHY it is. TRAIL exists because the participants listed above recognized that there was an incredible wealth of technical/historical content within academic library collections that was virtually inaccessible because the vast majority of it is not cataloged. TRAIL makes that content digital, discoverable, and openly accessible.

Supplemental Content: View/Download DeSart Poster

005. Dataset to DOI Using SWORD
Ann Devenish, Marine Biological Laboratory and the Woods Hole Oceanographic Institution

This poster describes the tools and processes by which datasets are deposited in an institutional repository (IR) and digital object identifiers (DOI) are acquired.

The Marine Biological Laboratory/Woods Hole Oceanographic Institution (MBLWHOI) Library and the Biological and Chemical Oceanography Data Management Office (BCO-DMO) have developed tools and processes to automate the ingestion of datasets and metadata from BCO-DMO for deposit into the Woods Hole Open Access Server (WHOAS) institutional repository. The system also incorporates functionality for BCO-DMO to request a DOI from the Library and streamlines the process by which XML files are generated for DOI deposit with CrossRef, a DOI registration agency.

This partnership allows the Library to work with a trusted data repository to ensure high quality data while the data repository utilizes library services and is assured of a permanent archive of the copy of the data extracted from the repository database. This research is being conducted by a team of librarians, data managers and scientists that are collaborating with representatives from the Scientific Committee on Oceanic Research (SCOR) and the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC).

Supplemental Content: View/Download Devenish Poster

006. Education Scholars’ Perceptions and Practices toward Open Access Publishing*
Lori Ellingford, Arizona State University | Sarah K. Brem, Arizona State University

Although OA publishing has been available since 1998, we know little regarding scholars’ perceptions of and motives for publishing in OA outlets. Among the disciplines using OA, Education has been relatively slow to adopt, yet the potential impact of OA could make OA one of the most important innovations in Education in the past few decades. We surveyed Education scholars using a 48-item survey. The survey contained subscales focusing on contemporary OA themes—issues identified through a comprehensive analysis of the major outlets for scholarly news in Education. Through open and axial coding, we extracted 9 themes, including ease of access and publishing, costs, support from colleagues and administrators, and perceived quality of OA outlets. (cont’d following page)
The sample comprised authors who were published between the years 2007-2010 at two OA journals based in the U.S. One hundred twenty-six individuals participated (56% response rate). The survey showed moderate to high reliability using Cronbach’s alpha. Among our significant findings: (1) Non-tenured faculty find it easier to use OA journals as opposed to print journals. (2) Non-tenured faculty believe it easier to get published in OA journals. (3) Scholars with 16+ publications and scholars at public universities find OA attractive because of low publication costs. (4) Compared to private institutions, scholars at public universities are more likely to read OA journals. (5) Scholars with less than 5 publications expressed hesitation to publish in OA journals.

We will discuss the implications of our results, focusing on how to improve perceptions and use of OA outlets.

**Supplemental Content:** View/Download Ellingford/Brem Poster

---

**007. Open Access Under the Radar: Embedding Discussion in Undergraduate Curriculum**

Isaac Gilman, *Pacific University*

Many attempts at open access advocacy are easily avoided by uninterested students and faculty—workshops, speakers, presentations at faculty meetings, or other one-off events. However, embedding investigation of open access into relevant curriculum ensures that students will have to engage with the issues—and will hopefully become interested and informed advocates for open access. At Pacific University, library faculty currently offer a 2-credit course on scholarly journal publishing for undergraduate students. Through lectures, discussions and assignments (including in-class debates) related to scholarly publishing, students have the opportunity to explore the economics, politics and practical considerations of open access to the scholarly literature.

After two years as a pilot, the course has recently been approved as a permanent annual offering in the University’s course catalog, and will be a core offering in a new undergraduate minor in editing and publishing that has been collaboratively developed by Library and English Department faculty. As it develops further, the course is also intended to provide training for student assistants who are employed by the Library to support its repository and publishing services—which serve to further advance discussion of open access on campus.

**Supplemental Content:** View/Download Gilman Poster

---

**008. GKR: Cooperative Repository Services**

Marlee Givens, *Georgia Institute of Technology*

In 2009, the Institute of Museum and Library Services (IMLS) awarded the Georgia Institute of Technology a 3-year grant to initiate the GKR statewide repository service in Georgia. The GKR's merit lies in its replicable practices and technology solutions for inter-institutional collaborations supporting scholarly communication. The GKR service provides: 1) hosting of independent DSpace instances for GKR participating institutions; 2) a central DSpace repository of harvested metadata from existing and hosted institutional repositories; 3) IR-related services that include guidance and training on metadata, content submission and rights management, digital preservation, and content digitization; and 4) a new, open source repository collection mapping tool to create a common discipline-based taxonomy across repositories with dissimilar academic and research vocabularies. With this open source tool, the GKR addresses a central challenge for statewide repositories -- joining content from partner institutions into a common system, which users can both browse and search centrally. The GKR mapping tool allows partners to map entire repository collections to discipline-based collections in the central repository using just a Web browser. The resulting mapping data is used when ingesting partners’ metadata to bring together their disparate content under a common taxonomy. In August, the GKR service is hosting a symposium on statewide and consortial repositories for other states and consortia considering the establishment of IR services, and a hands-on workshop for those interested in learning more about Georgia's approach to developing cooperative repository services.

**009. “Google, and Twitter, and Bitly—oh my!”: Novel Approaches to Repository Collection Development**

Robert Hilliker, *Columbia University*

In order to capitalize on the recent increase in traffic to Academic Commons, the digital research repository of Columbia University and its affiliates, we have implemented a couple of new collection development strategies that focus on faculty who are strongly motivated to enhance their visibility on the Web.

The first is using a set of weekly Google Alerts and RSS feeds to track mentions of Columbia researchers in the news media. We then contact them with a customized outreach message inviting them to deposit their work. We also tweet about new content added in this way.

The second is emailing Columbia researchers a monthly usage report for their content detailing how many times each of their works in the repository has been viewed and downloaded, both that past month, and over their lifetime. This report also includes a link to our newly
designed self-deposit form, thereby providing a gentle, but regular, reminder to faculty about the value of adding their work to the repository. The third is to further enhance the visibility of Academic Commons content by faculty in the news by coining a Bit.ly URL that points to a dynamically-generated search results page for their content and then using Twitter to push that URL out to relevant Twitter feeds, including the faculty member’s and those of the department or center with which they are affiliated. We also use hashtags to take advantage of the additional boost to visibility that they can provide.

Together these efforts have proved highly successful at motivating faculty to deposit their work in an ongoing fashion: every week we identify a dozen or more researchers and anywhere from a third to half respond by depositing work in the repository; every month over 1,000 email stats reports go out and, as a result, dozens of new items are deposited; several times a week we tweet about this content, generating hundreds of additional pageviews and downloads every month.

**Supplemental Content:** [View/Download Hilliker Poster]

### 010. VIVO: A Partner for Open Discovery
Kristi L. Holmes, Washington University in St. Louis School of Medicine

VIVO is an open source semantic web platform that enables the discovery of research and scholarship across disciplinary and administrative boundaries through interlinked profiles of people and other research-related information. VIVO is populated with information about researchers – allowing them to highlight areas of expertise, display academic credentials, visualize academic and social networks and display information such as publications, grants, teaching, service, and more. Profile information can be programatically imported from authoritative data sources such as institutional records and bibliographic and grant databases. Moreover, VIVO provides a way to highlight the efforts of an organization’s researchers in Open Access journals and offers institutions an elegant way to highlight their institutional repository as a critical campus resource.

By storing data in RDF and using standard ontologies, the information in VIVO can either be displayed in a human-readable web page or delivered to other systems as RDF, allowing the open researcher data in VIVO to be harvested, aggregated, and integrated into the Linked Open Data cloud. Data reuse and repurposing enhances the accuracy and currency of the data while increasing efficiencies across an organization. VIVO and other compatible applications produce a rich network of information across institutions, organizations, and agencies that can be searched to foster collaboration and enable open research discovery. The VIVO software and ontology are publicly available at http://vivo.sourceforge.net along with content that supports implementation, adoption, and development efforts around the world. To learn more about VIVO, please visit http://vivoweb.org/. VIVO is supported by NIH Award U24 RR029822.

**Supplemental Content:** [View/Download Holmes Poster]

### 011. Metadata for the Lifecycle of Research Data
Rhonda Marker, Rutgers University

There are many ways of depicting the life cycle of research data, and this is just one of them. I’ve chosen to show some of the complex objects that are created at many different point of that cycle. But what I want to focus on especially is the role of metadata.

At the Rutgers University Libraries, we are prepared to manage the full life cycle of research data and its related components. We are developing a Research Data Metadata Application Profile that can be applied to the data itself, to the research context, and to the objects associated with the data. The few metadata elements and vocabularies depicted here are meant to challenge your imagination and encourage discussion. It certainly has been a challenge to us! The file formats alone test even the most experienced data librarian! We have applied this metadata to the several research data projects archived in our institutional repository, RUcore.

In the process of identifying metadata elements and vocabularies for data, we are developing vocabularies that we believe are common to data objects and the research process in a wide variety of contexts. We recognize the need for collaboration to develop shared best practices and in some cases even standard core element sets and vocabularies.

We don’t have all the answers, but we begin with these questions:

- How can we best develop these shared guidelines and practices?
- How do we involve the appropriate stakeholders in this process?

**Supplemental Content:** [View/Download Marker Poster]
013. IR Workflow Management System: Using Web of Science and RefWorks
Jenny Oleen, Kansas State University

A major component of OA outreach at Kansas State University is using updates from Web of Science to identify recent articles authored by KSU faculty. We contact authors individually via e-mail and invite them to deposit the article in our IR, the K-State Research Exchange. By contacting authors when the article is first published, we have greater success in obtaining their final manuscript, if required for deposit.

The Library does the entire deposit on behalf of faculty and uses an account from the university-wide RefWorks license to manage the deposit workflow. Refworks has several features that lend itself to this role, including

- User defined fields for adding notes. We use these fields to note the author's email address, if an author has been contacted, if we're waiting for reply, etc.
- The ability to group items into folders, which facilitates sharing the workload among several staff. For example, those items for which we need to verify publisher policies can be placed in one folder; items that are ready to deposit can be placed in another folder.
- The ability to attach a file, such as the author's manuscript, to a record in RefWorks.

This approach has enabled us to increase the number of items deposited, cut down on processing time, and most importantly, interest a greater number of faculty in depositing their work. Although we have used Web of Science and RefWorks, most institutions have access to similar journal databases and bibliographic software (EndNote, Zotero) that can serve this purpose.

Supplemental Content: View/Download Oleen Poster

014. OpenSky: Linking Open Scholarship with Equipment and Services
Jennifer Phillips, National Center for Atmospheric Research

OpenSky is the open access institutional repository for the National Center for Atmospheric Research (NCAR). OpenSky supports the atmospheric science community by providing free and open access to NCAR research, increasing the visibility and impact of that research. OpenSky also supports the University Corporation for Atmospheric Research (UCAR), the management entity for NCAR, by providing metrics for internal reporting requirements. OpenSky contains a variety of materials, including peer-reviewed, published articles; proceedings, presentations, and posters; NCAR technical reports; and archival materials. While OpenSky is in many ways a typical open access repository, in addition to cataloging and serving documents, we are tracking the relationships between published scholarship in our repository and the UCAR assets and services used to create that scholarship. For example, a significant amount of geoscience scholarship relies on the use of NCAR's supercomputing facilities. By tracing the relationship between
computing time and the resulting publications, we begin to be able to trace the impact that supercomputing facilities have on the scientific community. We have developed controlled vocabularies for OpenSky that will allow us to manage such linkages between citations for publications and the UCAR services and equipment they used. Another related initiative is to link UCAR-managed datasets to publications in OpenSky by working with UCAR data centers to assign Digital Object Identifiers (DOIs) to datasets and create web-based dataset-to-publication linkages. Through these initiatives, OpenSky manages not only the intellectual output of our scientists and engineers, but also the relationships between organizational assets and the broader scholarly record.

Jason Priem, University of North Carolina at Chapel Hill

Citation studies have established OA’s value for enhancing traditional impact. However, until now it’s been difficult or impossible to measure the presumed OA impact on clinical practice, public awareness, and informal scholarly discussion. This is changing, though; the discussion and use of scholarly products is moving online, giving us a chance to measure these once-hidden impacts.

Total-impact (http://total-impact.org) is a free, open-source Web application that makes these “invisible impacts” visible. Users input research products—from datasets to articles to code repositories—and get back a report showing discussion, uses, and downloads of these products across Wikipedia, Twitter, Facebook, Mendeley, and many other Web environments. These reports are produced in real time and can be shared, downloaded, or embedded in a blog or publishing platform.

We’re excited about total-impact because it makes this data freely available to all, and can support a reward system in which decision-makers look beyond the Impact Factor to promote accessibility, engagement, and a broader understanding of “impact.” Ultimately, tools like total-impact could help support a new ecosystem of scholarly communication in which evaluation becomes a more collaborative, open, crowdsourced endeavour.

016. Open Content and the BHL: Breathing New Life into Old Data
Trish Rose-Sandler, Missouri Botanical Garden

The Biodiversity Heritage Library (BHL) is a consortium of natural history and botanical libraries that cooperate to digitize and make accessible the legacy literature of biodiversity held in their collections and to make that literature available for open access and responsible use as a part of a global “biodiversity commons.” The primary principles of the BHL are to be Extensive, Open, and Global. We’ll discuss how the principle of Open has driven decisions to provide historic content in as many forms as possible for others to harvest and reuse. We’ll highlight some creative uses of BHL data by other projects and individuals and finally we’ll discuss some factors that we believe are key to making open content more successful.

Supplemental Content: View/Download Rose-Sandler Poster

017. Promoting Open Access Resources Beyond Faculty: BYU’s Scholarly Communication Services
Elizabeth Smart, Brigham Young University

Brigham Young University’s Harold B. Lee Library (HBLL) strives to support faculty in all areas of scholarly communication by hosting a range of tools to help them improve the visibility of their scholarly work, make their work accessible to other researchers, and manage digital scholarship. Specific resources focused on these goals include ScholarsArchive (an open access institutional repository; http://lib.byu.edu/sites/scholarsarchive/) and the Scholarly Periodicals Center (digital publishing services; https://ojs.lib.byu.edu/spc/index.php/index).

Teaching faculty have traditionally been the sole target for promotion of HBLL’s scholarly communication services. More recently, HBLL has expanded its efforts to new audiences on campus, including public relations specialists, department and college assistants, and student publishing advocates. Like the library, these campus professionals are also tasked with supporting the scholarly and creative work of faculty and students, and with making the products of faculty work more visible. For example, some academic colleges employ public relations specialists who are eager to increase the visibility of faculty research to the media and college alumni. Institutional repositories are useful tools for public relations professionals because IRs provide free access to faculty research for the media, alumni and other interested readers. HBLL has also been working with advocates for student publishing who are seeking sustainable ways to feature student work. Library-hosted open repositories and digital publishing services help these advocates as they mentor students in producing student journals and scholarly conferences. (cont’d following page)
Microsoft Academic Search (MAS) is a free service used by scholars, scientists, students, or anyone with a research question. There are currently nearly 40 million scholarly papers indexed within MAS, with tens of millions more in the process of being indexed. MAS is an open tool. We have developed an API that allows all of the information in the service to be accessed and reused for noncommercial purposes. We want the community to build compelling tools and experiences on top of our indexed data. The API program is starting to emerge. One example is the Eigenfactor Project, out of the University of Washington, which is using the MAS API to harvest article data and develop novel methods for evaluating a paper’s impact. Another example is ScienceCard, which integrates an author’s publication list with Twitter. Also of note, Microsoft Academic Search has generated nearly 20 million author profiles from the data we index. These profiles may be easily embedded in outside sites. Universities can generate code for all their researchers indexed within MAS and embed these profiles within an institutional repository.

Patron-Driven Institutional Repository Development
Donald Taylor, Simon Fraser University

In August 2011 the Simon Fraser University Library launched a new IR, Summit, based on the open source Drupal content management system. Prior to this, IR Services had been very low profile due to the perceived lack of flexibility of the current IR platform, lack of extensive Java programming skills and our inability to accommodate enhancement requests from faculty.

Moving to Drupal allows the SFU Library to take advantage of the global pool of Drupal developers and our in-house Drupal expertise. For example, only three modules were written from scratch while the other seven required modules were modified ones already developed by Drupal developers. As a consequence of this change of platform, we now have a dynamic and flexible platform that allows us to collaborate with faculty on enhancements to develop more responsive IR services. Faculty requests for enhancements drive IR development: examples include small tweaks like adding required HTML metadata headers to ensure indexing by Google Scholar and export of records to Zotero bibliographic software; modifications of author pages; giving users the ability to version documents; restricted collections for purposes of document sharing and pre-publication work; record display enhancements; search result enhancements; ongoing development of ways to synch collection content in the IR with departmental websites. These enhancements help to bring in new users.

As well, the SFU Library’s new Scholarly Digitization Fund, where the Library pays for digitization of faculty research, has brought in non-traditional IR users; provided a further opportunity to promote OA; and broadened the Library’s scope of what belongs in an IR.

**Editor’s note:** Abstracts with an asterisk were accepted, but not presented, at the SPARC conference