BEYOND BONECOMMONS

RECENT DEVELOPMENTS IN ZOOARCHAEOLOGICAL DATA SHARING

Sarah Whitcher Kansa and Eric C. Kansa

Sarah Whitcher Kansa is Executive Director of the Alexandria Archive Institute and Editor of OpenContext.org in San Francisco, CA, USA.
Eric C. Kansa is Lead Developer of Open Context at the Alexandria Archive Institute and a Researcher at the University of California, Berkeley.

BoneCommons (http://alexandriaarchive.org/bonecommons/) was developed in 2006 by the Alexandria Archive Institute (http://alexandriaarchive.org/) as an open access system to advance communication and sharing of scholarly content within the global zooarchaeological community. It works as a community hub to enable sharing and discovery of images, conference presentations, and papers.

Because web-based dissemination costs are now so low, social media[1] systems generally permit a great range and diversity of content to be shared. One can just as easily post an image, a 140-character "tweet" or an entire thesis. Social media technologies such as blogging, podcasting, and wikis lower barriers for individuals to reach a global audience. These same low barriers also enable social media to be more conversational because commentary and continued conversation are easily supported (see also Morris, this issue).

However, the potential removal of the "gatekeeper" (i.e., editorial oversight of print publication) means that any and all content can be shared. Thus user-generated content can run the gamut from distracting commentary to valuable research content. This openness and ability to rapidly publish anything sometimes lies in tension with professional expectations for quality and vetting of material. Any attempt to apply social media in scholarly contexts must confront these issues, and balance the affordances of easy, rapid dissemination with the imperative of maintaining professionalism and quality (see also Harley et al. 2010). BoneCommons represents our attempt to adapt social media to scholarly communications in zooarchaeology.

Stepping Back: What do Users Need?

The original BoneCommons was a forum aimed at facilitating discussion and contact between zooarchaeologists worldwide by offering a place to post papers, ideas, images, questions, and comments. As more content was added to the site, browsing and searching became challenging, impacting the site's mission of providing "easy access to quality content." This, together with spam and security issues made the site too costly to maintain on a volunteer basis. Informed by new social media technologies, a better understanding of what users need, and lessons learned about "under the hood" requirements for professionally oriented web dissemination, we undertook a major remodel of BoneCommons in 2009.

This remodel was timely, as we were engaged in a two-year study[2] exploring how open technologies can best meet the needs of the diverse communities of scholars working with digital archaeological content. While the study's focus on primary analytic data prompted significant changes to Open-Context (http://opencontext.org), an open access data publishing system that we developed and maintain, some aspects of it also directly informed our reconfiguration of BoneCommons.

One of the key outcomes was the importance of linking primary analytic data with other datasets and scholarly content. Analytic data is created as part of broader research programs that typically result in published scholarly syntheses. These syntheses can both be enriched by access to primary data and help make primary data themselves more intelligible and meaningful. Our study highlighted the need to reference and cite across these various forms of content (which, on the Web, involves hyperlinks). We also noted that users demand efficient search and discovery functions, as well as the ability to easily export data from one application to another.

With these needs in mind, we remodeled BoneCommons in 2009 with new software that offered better functionality, design, and security. The new BoneCommons is built with free, open source content management software offered by Omeka (http://omeka.org/), an ideal solution for a system with little funding and outside technical support, and offers the following features:

- Clear citation and stable linking for every item.
- Standard Dublin Core library metadata and digital library protocols to facilitate archiving.
- A contribution form for users to upload their own content and relevant metatdata ("information about information").
- A clean and professional appearance (as opposed to the forum-style of the old site).

In BoneCommons, all contributions are vetted by the site's editor. This oversight does not substitute for peer-review, since BoneCommons is not intended to replace journals or conferences. Rather, BoneCommons can be more relaxed in reviewing content because much of the material posted in BoneCommons was vetted elsewhere (by conference review committees or publishers), and BoneCommons' review process mainly addresses relevance and completeness of metadata, especially around citation. In cases where researchers post images to BoneCommons in order to seek help with identifications, vetting need not be too onerous. Thus, we have tried to retain social media's efficiencies for rapid publishing while adapting social media to better meet professional needs, especially with regard to citation and editorial oversight.

Monitoring user searches in BoneCommons helps us address search and discovery needs. Figure 1 shows that searches monitored over three months resulted in a classic power law graph, where a few topics dominate and the rest fall into a "long tail" of topics that cannot be categorized. Nearly half the searches requested people or publications, while the remaining 51 percent of searches used more idiosyncratic terms. This long tail of search terms, each of which occurs only rarely, highlights the great breadth of user interests. These results help demonstrate the importance of continual improvement in search and discovery services, since researchers are often looking for very specific "needles" in increasingly large and complex "haystacks."

Adapting BoneCommons to better meet user needs has helped improve the site's impact. At the same time, professional acceptance of web-based scholarly communications is increasing in zooarchaeology. After the 2006 ICAZ international conference, abstracts of conference papers and posters were posted to BoneCommons. Presenters were given the option of posting their communications on BoneCommons, but only a handful chose to share in this way, opting instead to wait years for print publication. By the time of the next ICAZ international conference (August 2010), the community's perception of sharing research openly on the Web had clearly changed. Of 700 oral and poster communications, over 120 were posted to BoneCommons[3] in the four weeks following the conference.

Many scholars now embrace the medium as a way to communicate their work in one way or another, whether it takes the form of sharing PDFs of publications, commenting on another researcher's work, or responding to a question. The change in BoneCommons use took some years to occur, but it is important to note because it demonstrates that the perceived barriers to adopting technologies in the "static" world of academia are not insurmountable.

Beyond BoneCommons

The variability of research interests and the multidisciplinary nature of scholarship highlight the need for scholars to draw upon a diversity of information sources and services. Our challenge is to find ways to facilitate use of distributed resources by the research community. We are experimenting with various ways of making BoneCommons function in a landscape of distributed information sources and services. One example of how web resources can be linked to facilitate information access is the relationship between BoneCommons and ZOOARCH,[4] an email list dedicated to zooarchaeology-related discussion. ZOOARCH prohibits attachments, so subscribers often post images on BoneCommons and share the link on ZOOARCH. Connecting the two resources in this way facilitates discovery and discussion across user communities (see Figure 2).

BoneCommons sees more active participation from its links with conversations on ZOOARCH. When postings come to users by email, they are more directly tied into everyday workflows. Furthermore, Figure 2 demonstrates how scholarly expectations can compel people to use social media, where the response of the third researcher to "ask the expert" drew that person into the conversation with the expectation that he would reply out of perceived scholarly obligation.

Opening Information Silos

BoneCommons and ZOOARCH offer different, complementary services. Rather than attempting to create the perfect resource, we should instead be targeting our funding and technological know-how to *facilitating information flow between resources* so that synergies can emerge. Sadly, though many web resources exist, they tend to stand in relative isolation from each other; that is, you have to know about each one and visit them to see their contents. Thus, though the Web has allowed people to share vastly more information than before, much of it is trapped within "information silos."

Building for "mashups" (simple ways of combining data from different web sources) can help mitigate this problem
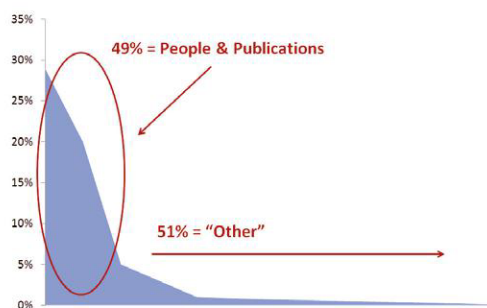

Figure 1: The "long tail" of 1132 searches leading to BoneCommons (searches from April – June, 2010).
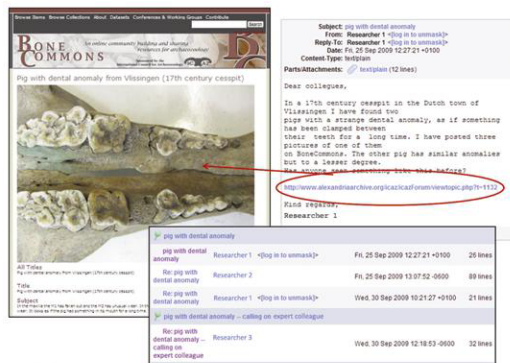
Figure 2: Diagram showing the flow of information between the ZOOARCH email list and BoneCommons. Counterclockwise from upper left: Researcher 1 posts an image and description on BoneCommons and then posts a question to the ZOOARCH list with a link to the image in BoneCommons. A discussion ensues on the ZOOARCH list between three researchers. In the end, Researcher 1 is pointed to a specific colleague who is an expert in the domain for further advice on the matter.

of isolated resources. For example, BoneCommons displays a feed from the ZOOARCH list's server, in essence drawing related content from one resource to another. This "-machine-readable" data facilitates web-based scholarly communications, allowing easy aggregation of relevant content so users of BoneCommons, for example, can immediately access content published by another source.

Machine-readable data can also help address an important need identified in our user-needs study—linking between primary analytic data and other research outputs, such as conference presentations and articles. Currently, we are experimenting with bringing relevant primary data to the BoneCommons community hub via Open Context's machine-readable data services to (see Figure 3). Open Context is an open access, web-based system that publishes editorially verified datasets, many of which are linked to print publications. It contains a variety of project datasets that span geographic regions, time-periods, and areas of specialization (Kansa 2010). Since only a portion of its content will be relevant to a given community, Open Context offers sophisticated services for querying machine-readable data. This allows BoneCommons to request a filtered subset of Open Context content so that only zooarchaeology-related data appears on the BoneCommons site. BoneCommons' "Datasets" page (Figure 3) draws relevant data from Open Context, demonstrating how a customized data "feed" can flow from one web resource to another so that relevant research content from multiple sources can be drawn together in useful ways.
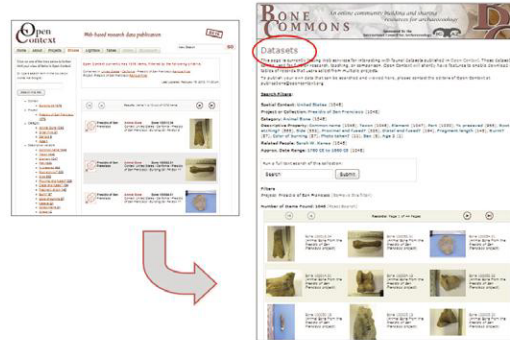

Figure 3: Custom feeds of zooarchaeology-related data flowing from Open Context into BoneCommons.

A Vision for Distributed and Collaborative Resources

Machine-readable data facilitates aggregation so that community hubs can offer rapid and efficient access to high-quality information from multiple, trusted web sources. Such aggregation benefits researchers facing tremendous time pressures. By allowing information to escape a given silo and enter a new context, machine-readable data helps bring relevant content to users rather than expecting users to go to that content.

By virtue of being based on the Omeka software application, BoneCommons itself publishes machine-readable data and is open for aggregation by other sites. Beyond the ZOOARCH list and Open Context, many other web resources can offer machine-readable data. Looking to the future, more widespread adoption of even very simple types of machine-readable data services such as the "Open Search" protocol can have a great impact for researchers. Open Search enables cross-site searching so relevant data can flow across different sites. Because few web resources will ever be able to achieve the scale and comprehensiveness necessary to meet the needs of highly specialized research interests, such services are needed to help users look beyond one site and efficiently search across many trusted and high-quality scholarly repositories. Such services can be useful building blocks for more sophisticated kinds of data integration and analyses pooling data from multiple sources.

Our experiments with BoneCommons demonstrate useful ways to adapt and deploy social media and machine-readable data to meet researcher needs. We look forward to seeing online reference collections and other data providers offer similar machine-readable data services. By making high-quality web resources that are open to reuse and aggregation, we can network zooarchaeological information services in ways that better reflect the vibrancy and collaborative spirit of the zooarchaeological community.

References Cited

Harley, Diane, Sophia Krzys Acord, Sarah Earl-Novell, Shannon Lawrence, and C. Judson King.
  2010 Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines. UC Berkeley: Center for Studies in Higher Education. Retrieved from: http://escholarship.org/uc/cshe_fsc (Accessed June 22, 2010).

Kansa, Eric C.
  2010 Open Context in Context: Cyberinfrastructure and Distributed Approaches to Publish and Preserve Archaeological Data. *The SAA Archaeological Record* 10(5):12–16.

Kaplan, Andreas M., and Michael Haenlein
  2010 Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons* 53(1):59–68.

Notes

    1. The term "social media" describes web-based applications that facilitate the creation and exchange of user-generated content (Kaplan Haenlein 2010), or publicly-available information published by end-users.

    2. This study was funded by a grant from the National Endowment for the Humanities and Institute for Museum and Library Services' *Advancing Knowledge: The IMLS/NEH Digital Partnership* program.

    3. http://alexandriaarchive.org/bonecommons/items/browse/tag/presentation+shared

    4. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=zooarch

Position Open: Post-doctoral Scholar (PS)
Location: Buffalo, New York

Post-doctoral Scholar Position 2011-2012, SUNY Buffalo. SUNY Buffalo seeks a Post-doctoral Scholar (PS) for its interdisciplinary Institute for European and Mediterranean Archaeology (IEMA). During a 10 month tenure, the PS teaches one graduate seminar (preferably on the topic of the symposium), organizes a symposium, and edits a subsequent volume reflecting IEMA's focus on post-Pleistocene European and Mediterranean anthropological and classical archaeology. Application letter, vitae, list of references, and 3-page description of proposed symposium topic, including intended invitees, must be received by March 1, 2011 for an August 2011 start, pending final budgetary approval. Email application or inquiries to: pbiehl@buffalo.edu. The University at Buffalo is an affirmative action/equal opportunity employer.

Throughout the session, it became clear that many of these initiatives share complementary insights and goals and have much to learn from each other as well as from similar initiatives in archaeology and museum practice.[2] Audience discussion highlighted the considerable variability in how zooarchaeological information is recorded, described, and archived (see also Driver 1992) as well as the need for academic and CRM-based incentives to make primary data accessible. Others noted how granting agencies and repositories are now requiring submission of all digital information and researcher expectations are shifting toward greater transparency and accessibility. As a wave of zooarchaeological researchers approach retirement, such changes are urgently needed as primary zooarchaeological data is highly vulnerable to loss and fragmentation. The projects presented here demonstrate how researchers are making serious collective efforts to improve communication and access in zooarchaeology and beyond.

References Cited

Driver, Jonathan C.
  1992 Identification, classification, and zooarchaeology. *Circaea* 9(1):35–47.

Notes

    1. ICAZ is a scholarly organization which hosts quadrennial conferences concerning zooarchaeology: http://www.alexandria archive.org/icaz/index.htm.

    2. A growing list of such resources can be found on the ICAZ website: http://www.alexandriaarchive.org/icaz/about_links.html.