

FORUM

Open Context, Data Sharing and Archaeology

An Open Context for Near Eastern Archaeology

Sarah Whitcher Kansa, Eric C. Kansa, and Jason M. Schultz

The common use by archaeologists of ubiquitous technologies such as computers and digital cameras means that archaeological research projects now produce huge amounts of diverse, digital documentation. However, while the technology is available to collect this documentation, we still largely lack community-accepted dissemination channels appropriate for such torrents of data. Open Context aims to help fill this gap by providing open access data publication services for archaeology. Open Context has a flexible and generalized technical architecture that can accommodate most archaeological datasets, despite the lack of common recording systems or other documentation standards. It includes a variety of tools to make data dissemination easier and more worthwhile. Authorship is clearly identified through citation tools, including web-based publication systems that enable individuals to upload their own data for review, and collaboration is facilitated through easy download and “tagging” features. Near Eastern archaeologists will benefit from Open Context’s flexibility to share a variety of content from diverse projects, no matter how large or small.

Is there a Future for the Past?

Simple lack of publication poses an under-recognized threat to cultural heritage preservation. Without sharing, irreplaceable knowledge of the past is one hard-drive crash away from oblivion. Such everyday data loss represents a tremendous ethical and professional failing in our discipline. Digital publishing can now help protect the rich and painstakingly compiled bodies of documentation developed in Near Eastern archaeology from the threat of loss. A recently released data-sharing web application, called Open Context (www.opencontext.org), enables researchers to publish structured data along with textual narratives and media (images, maps, drawings, videos) on the web. This new system, described in detail below, provides a cost-effective, scalable solution for many data-sharing needs in Near Eastern archaeology and related fields.

Documenting, Preserving, and Sharing the Past

Advances in digital technology are transforming Near Eastern archaeology. New tools for data collection, such as electronic distance measurement devices (EDMs), global positioning systems (GPS), digital cameras, and video recording, and the growing popularity of handheld data-entry devices, mean that the practice of archaeology increasingly results in “born digital” documentation. This is typically much richer and more comprehensive than traditional paper and photographic film recording techniques. The continuing decline in storage costs and the growing sophistication of database systems help fuel this drive for more complete and thorough field recording and documentation. Digital documentation, coupled with digital communication via the Internet, permits far more rapid and comprehensive dissemination of field research.

Besides making distribution highly cost-effective, the Internet is a powerful means to share large collections of rich media and complex data. These types of content are important components of both museum collections and excavation documentation. Many museums now display portions of their collections online and some research projects have online databases documenting their excavation and survey results. Çatalhöyük, Tel Halif, and other Near Eastern sites have a rich online presence. The CyArk 3-D Heritage Archive Network provides a searchable archive of free 3-D scans and maps of World Heritage sites. The pioneering Perseus Digital Library has a rich and ever growing collection of texts, images, and other media for classical studies and other areas, while the Cuneiform Digital Library makes an impressive collection of early Near Eastern texts openly accessible. The public is getting involved as well. For instance, the commercial photo-sharing site *Flickr* currently has over fifty thousand photos of items in the British Museum, contributed by public enthusiasts fascinated by the historical and aesthetic achievements of the past.

Data-Sharing Challenges in Archaeology

In spite of these recent advances, the current reality of archaeological data sharing is not living up to its full potential. While costs are in sharp decline, many scholars still lack the means to share their field research easily. Many have difficulty seeing incentives for sharing and remain unaware of larger trends toward more open and rich forms of scholarly publishing. Thus, the dissemination of archaeological research remains a challenge to many in our field.

Among the primary technical and conceptual issues in sharing field data is the question of how to codify our

documentation. Archaeologists generally lack consensus on standards of recording and tend to make their own customized databases to suit the needs of their individual research agendas, theoretical perspectives, and time and budgetary constraints (see also Denning 2003; Hodder 1999). Because of this variability, databases need extensive documentation for others to decipher their contents. This type of documentation is often called “metadata,” a term that is typically defined as “information about information.” Metadata, such as titles, keywords, author, and catalogue numbers, enable library users to find relevant publications. Likewise, metadata documentation associated with archaeological datasets can help others find and decode those data. However, adding useful metadata to content typically requires time and expertise, thereby deterring many from sharing.

Even if we find solutions to documenting the diversity of archaeological content, the size and complexity of archaeological databases create challenges that even expert metadata documentation cannot solve. Large archaeological databases often include hundreds of thousands of individual records created by multidisciplinary teams, all in complex relationships. If a dataset needs to be downloaded and deployed on appropriate software, it will still be very difficult to use even with adequate documentation. Once it is deployed, users will have to familiarize themselves with a project’s database organization and interface. The steps involved in downloading and deploying such databases require too much effort for casual browsing and searching. Thus, making datasets available for download (even with adequate metadata) is not an ideal solution for archaeological communication if the data are not easily “digestible” by others.

A more ideal solution is to serve archaeological databases in dynamic, online websites, thus making content easy to browse and explore. Unfortunately, this typically requires complex and expensive custom web development. Thus, only a handful of very-well-funded projects offer access to databases of primary results via the Internet. The enormous and incredibly rich Çatalhöyük database represents just this kind of project-specific data sharing. Its extensive catalogue of excavated contexts and finds facilitates analysis and collaboration among the project’s large team of specialists. While this is a fundamental contribution to scholarship, Çatalhöyük’s system is not readily scalable. If other projects seek to adopt Çatalhöyük’s online database to share their own content, they would have to conform to its recording system.

Most Near Eastern archaeological projects take place in smaller research programs with less funding and technical support than Çatalhöyük. These smaller projects have little capacity to develop their own customized, web-accessible database solutions. They may develop rich bodies of documentation, but without Internet dissemination much of this material will never

see publication simply because this vast amount of content cannot be accommodated by print publication. The paper format is simply not up to the task. Therefore, the thousands of bones, seeds, potsherds, lithics, and other artifacts and ecofacts that are analyzed and recorded, as well as the maps, photos, and log entries associated with a typical project, almost never see publication beyond summarized forms.

Open Context: A Data-Sharing System for Near Eastern Archaeology

Sharing complex excavation and museum collection databases represents a new set of technical, conceptual, and incentive problems. There has been great progress on many of the technical and conceptual problems involved in pooling and integrating the complex and un-standardized data generated by researchers. ETANA-DL, led by James W. Flanagan and digital library pioneer Edward Fox, has successfully demonstrated a framework for interoperability and integrated search, browse, and analysis tools for several Near Eastern excavation datasets (Ravindranathan *et al.* 2004).

Research in data sharing has continued, and has led to the development of new systems now in demonstration. With financial support from the William and Flora Hewlett Foundation, the Alexandria Archive Institute (AAI) developed Open Context to help make the dissemination of cultural heritage collections easier and more cost effective. Open Context is a highly flexible database that enables researchers to publish structured data along with textual narratives and media (images, maps, drawings, videos) on the web. It provides an easy to use, yet powerful, framework for publishing, exploring, searching, and analyzing multiple museum collections and field-research datasets. The system draws on powerful, open-source technologies (MySQL, PHP, and Dojo AJAX), making Open



The structured data that researchers publish with Open Context can communicate with and draw from other systems.

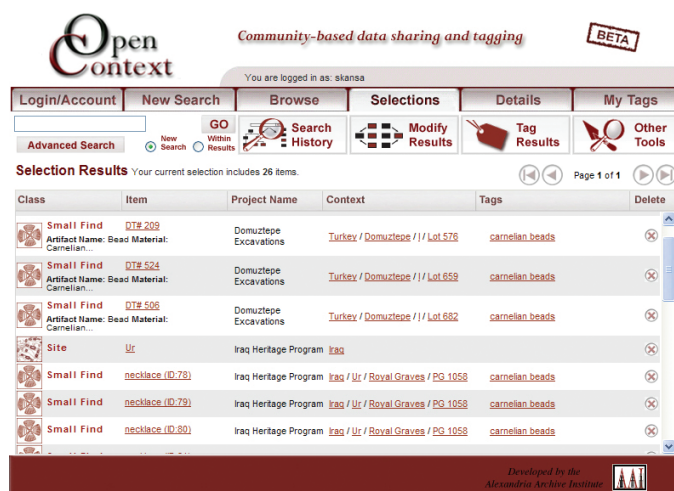
Context easy to implement even for organizations without their own servers. These technologies are also widely accessible and supported by a large developer community. Thus, Open Context can be easily deployed, maintained, and modified by anyone with standard web-development skills.

Open Context now supports a wide variety of cultural heritage applications, including field research and museum collections. Most recently, the system imported over 120,000 items from Brown University's excavations at the Great Temple of Petra, led by Martha Sharp Joukowsky. This flexibility stems from over twenty years of development and field-testing of database designs by David Schloen, head of the

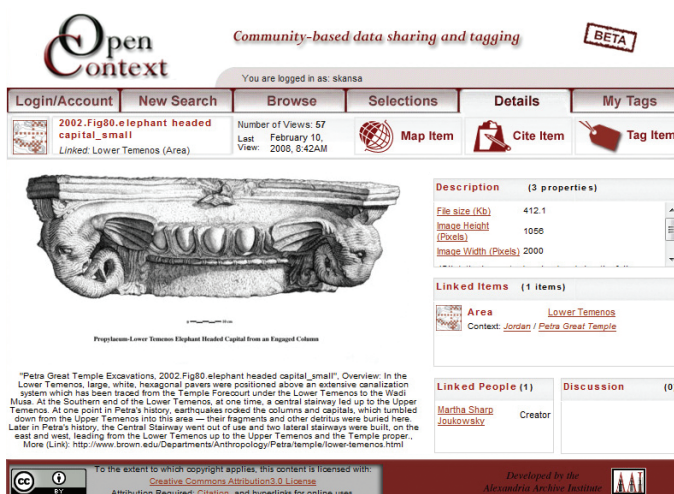
University of Chicago OCHRE ("Online Cultural Heritage Research Environment") system (Schloen 2001). While OCHRE provides sophisticated data-management tools targeted for active research projects, Open Context uses a subset of the OCHRE data structure (ArchaeoML) to support streamlined, web-based access and community organization of diverse cultural heritage content. Schloen designed the ArchaeoML structure to accommodate cultural heritage datasets without imposing rigid predetermined standard vocabularies or recording systems. Overly rigid standards may inhibit innovation in research design and poorly accommodate "legacy" datasets (Kansa 2005). The flexibility of ArchaeoML enables Open Context to deliver content from many different research projects and collections. A web-based publishing tool called "Penelope" enables individual contributors to upload their own data tables and media files and submit them for review and publication in Open Context. This tool enables web publication of research while ensuring that a project's original recording system and terminology are retained.

Citation, Access, and Copyright

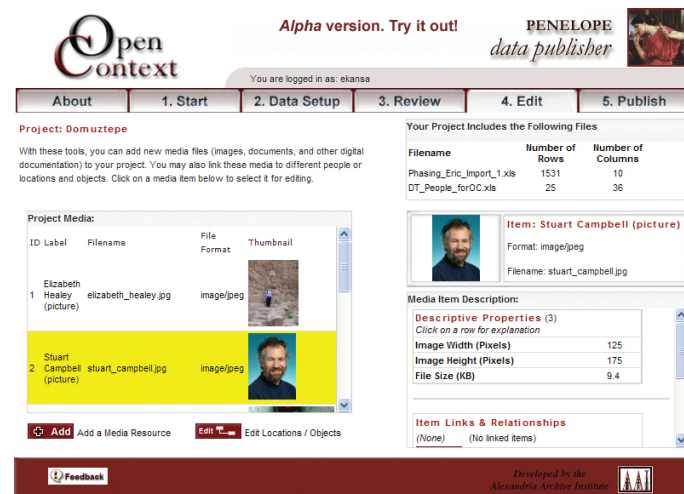
While Open Context and other technologies for cost-effective Internet publication of archaeological data are emerging, several important social, legal, and incentive issues remain. Data hoarding, sloppiness in record keeping, and the lack of positive rewards for data publication inhibit many researchers from participating in online scholarship (Willinsky 2006:21). While the purpose of this article is to introduce a data-sharing framework and not to detail all of the complexities of copyright, some discussion is warranted (and should be the subject of future discussions).



Results of a search for "carnelian" in Open Context, showing items from multiple projects.



An image in Open Context linked with its small finds registry record and context.



A view of /Penelope/, Open Context's web application for data publication.

Briefly, US copyright law makes a distinction between public domain “facts” and copyright protected “expressions.” Most archaeological datasets contain both factual data (tabular analytic data such as measurements and species identifications) and expressive content with some degree of authorial or creative originality (free-form notes, drawings, images). Because archaeological documentation usually mixes fact and expression (in the legal sense), and because the threshold for determining copyright originality is generally low, copyright will typically apply to most archaeological field data (see Kansa, Schultz, and Bissell 2005). Thus, measures to manage copyright must be built into archaeological data-sharing systems. Open Context takes its copyright policy cues from similar “open data” initiatives (such as Science Commons, PubChem and Freebase) and policy moves by the likes of the National Science Foundation (National Science Foundation 2007:29) and the Scholarly Publishing and Academic Resources Coalition (SPARC).

Open Context contributors retain copyright to their own content. This policy encourages researchers to publish their content in multiple venues (including journals and books). However, current copyright standards restrict all copying of works without explicit permission from the copyright holder. This complicates digital preservation, because many preservation programs copy data for storage in different repositories across the Internet. Such redundancy offers an important safeguard against the failure of any one repository. In addition, the legal constraints of copyrighting make it more difficult to build upon the research of others.

Legal copyright is not the only, or even the most important, form of protection in scholarship. When one publishes factual data even in a paper journal or a conference presentation, that factual data is not protected by copyright. However, professional norms expect citation of the creators of factual data. Limited transparency and unequal power relationships in the academy help fuel fears of being “scooped” by mentors and colleagues. Without transparency, such cases of academic misconduct result in “he-said, she-said” stand-offs. However, public exposure may be the best deterrent against unethical behavior. The Near Eastern archaeological community is small, and there are relatively few recognized publication venues. Open Context timestamps the accession of new collections and clearly identifies authorship, and all of this information is quickly reached via Google. Enhanced access and exposure can make it easier for the community to monitor itself and guard against intellectual dishonesty in professionally recognized journals. As a positive incentive, Open Context’s radical transparency in excavation data could enhance the credibility of related published

interpretations, an effect already noted in biomedicine (see Piwowar, Toger, and Fridsma 2007).

Therefore, Open Context, like many other open access (OA) systems, has adopted policies requiring use of standard copyright licenses that give explicit permissions for reproducing content. The nonprofit organization Creative Commons has developed the most-widely adopted set of copyright licenses used by OA systems. All Creative Commons licenses permit copying and reuse of content (with certain optional conditions and restrictions) on the condition that the copyright holder is properly attributed (Brown 2003). The Massachusetts Institute of Technology, Rice University, and many other scholarly programs publish material under Creative Commons’ licenses, as do a growing number of professional journals. This legal openness ensures that Open Context content can be easily shared and reused in both instructional and research applications.

Open Context content can be moved freely into other data systems, translated into other languages, or incorporated into books or course packs as long as contributors receive proper citation. By removing the legal barriers to sharing, research and instruction become easier. These licenses also benefit data authors, because they gain greater exposure and wider attribution as their contributions see new uses. The OA and Creative Commons licensed journal PLoS Biology is extremely competitive and has an impact factor rivaling established leaders such as Nature and Science. To guide users with proper attribution, Open Context automatically generates citation information and a stable URL for each item in the database. Reproduction of an item must link back to Open Context, and these links are counted and weighed by commercial search-engines (especially Google). Thus, individuals and their content gain higher search rankings.

The screenshot shows the Open Context website interface. At the top, the logo and tagline "Community-based data sharing and tagging" are visible, along with a "BETA" badge. A user is logged in as "ekansa". The main navigation bar includes tabs for "Login/Account", "New Search", "Browse", "Selections", "Details", and "My Tags". The "Details" tab is active, showing information for "Item: Locus 20". The page includes sections for "Context (click to view)", "Description", "Phase", "Phase Description", "Stage", "Item Notes", "Contents", "Linked Items", "Linked Media", "Linked People", and "Proper citation for this item". The citation information is automatically generated for each item.

Citation information is automatically generated for each item in Open Context.

Longevity support from the Internet Archive means that data will be retrievable even if Open Context itself goes off line. Bibliographic information stored in Open Context is also expressed in a standard that is readable by the open source Zotero (<http://www.zotero.org>) citation management tool. Zotero enables researchers to capture bibliographic information automatically when they use internet-based library and digital repositories, including Open Context. These citation features and copyright policies help encourage data longevity, access, and reuse. Thus, they align with established ethical guidelines advocated by the Society for American Archaeology (1996), especially Intellectual Property (Principle #5), Public Reporting and Publication (Principle #6), and Records and Preservation (Principle #7).

Open Context is an alternative to all-rights-reserved copyright and all-data-withheld research; it offers a different path, with different risks and opportunities. Digital communications make copyright violations easier, and copyright laws are commonly violated online. However, as witnessed by the Web's exponential growth and increasing use for scholarly communication, many judge opportunities to be greater than risks. Furthermore, copyright owners are not required to be driven by commercial gain! Many authors and artists desire non-economic benefits such as reputation and contribution to the advancement of scholarship over sales of content (and in the case of "raw data," commercial potentials seem minimal). Thus, we believe the benefits of enhanced exposure offered by open scholarly publishing helps the careers of typical scholars more than defensive attempts at copyright control.

Web 2.0 Tools for Near Eastern Archaeology

Open Context's interface offers straightforward browse, search, and analysis functions. Users have a variety of options to find materials in Open Context, including simple "Google-like" text searches and more sophisticated, advanced searches that use complex query logic. Simple charting tools help with data visualization, and a selected dataset (potentially drawing records from multiple projects) can be exported into common formats, such as Excel.

To help make sense of this widely varying documentation generated by archaeological research, Open Context has a variety of "Web 2.0" tools and features (that is, features that facilitate collaboration and sharing). To link database content with written narratives, Open Context automatically generates reciprocal hyperlinks with weblog or other publishing systems, such as e-journals that support the "ping-back" standard. Scholarly "bloggers" can therefore reference and discuss an item or a set of items from Open Context,

and it will automatically link to that discussion. This adds value to field research and material collections because it integrates that content with active discussions and debate. In addition, Open Context uses a folksonomy system to try to cultivate some of the same "peer-production" systems so successful in building rich bodies of knowledge in other online communities (Benkler 2006:75–81). Folksonomies are cost-effective and simple tools that enable a community of users to add value to pooled content by identifying and annotating items of interest. Users can "tag" items in Open Context with common keywords and phrases, and thereby establish and share meaningful links between items from different projects and collections, even if these projects use different recording systems. Open Context enables users to tag items either individually or collectively (i.e., users can assign a tag to items in a query result set). When query result sets are tagged, the history of query composition is automatically linked to the tagging event. Users can also further annotate and explain the rationale behind their tag assignments. Tags can be used to save search selections for future reference and to share sets of items with colleagues.

Recent experiments suggest folksonomy systems offer annotations of sufficient quality to meet some needs of museum professionals (Bearman and Trant 2005; Trant 2006). Currently, Open Context documents the authorship of each tagging event, and users can filter out tags and tag authors they consider to be unreliable. The system will soon be enhanced to recognize scholarly credentials and authority better. We also aim to provide the option for users to apply professionally developed standard vocabularies, such as the Getty Art and Architecture Thesaurus or other more sophisticated data integration systems.

Benefits to Near Eastern Scholarship

One of the greatest benefits to Near Eastern scholarship is that Open Context acts as a (near) "universal recipient" for the vast variety of information generated in our field. Open Context can accept and integrate diverse content from a huge project, such as ten years of excavation drawings, photos, small-finds databases, maps, and specialist analyses at Domuztepe, a twenty-hectare Neolithic site in southern Turkey. It can also house very small projects that might consist of only a few field notes and photographs. Without a simple and free means of publication, these small or incomplete projects are at risk of loss. Thus, Open Context provides a very economical model for transparency and greater access to a vast variety of Near Eastern archaeological research, past and present, large and small. Data contributors can choose when to publish with Open Context, long before, during, or after print publication.

Many of the datasets in Open Context may best function as elaborate appendices to supplement print publications. With sufficient community contributions, feedback, and support, Open Context will expedite reference searches and facilitate reuse of content. This pooling of primary data, in a system that makes content easy to access and reuse, enables broad regional syntheses that are more comprehensive and more analytically rigorous than are currently feasible (Kansa 2005; Kintigh 2006).

Open Context in the Larger Context

People connected to the Internet seem to have an insatiable appetite for even the most obscure information (Anderson 2006). If accredited scholars will not make their research openly available, others will meet this demand. "Peer production," the voluntary efforts of enthusiast communities, is rapidly becoming a major economic force, disrupting numerous software and publishing industries (Benkler 2006). According to current Alexa rankings, *Wikipedia*, the *ne plus ultra* example of peer production (developed, edited, and financially sustained entirely by volunteers), ranks in the top twenty most-visited websites on the Internet. Furthermore, the journal *Nature* recently demonstrated that the accuracy of the *Wikipedia* (in some subjects) is comparable to professionally edited encyclopedias (Giles 2005). This is not to claim that scholarly communication should be assessed only in terms of Internet popularity; instead, we call for greater engagement with other scholarly and public communities. OA publishing systems like Open Context can facilitate such engagement without sacrificing editorial control and professionalism. "Going online" does not mean abandoning peer-review and adopting the radical egalitarianism of *Wikipedia*. The peer-review process can work with the tremendous economic and impact advantages of open digital dissemination.

Critics who question the financial sustainability of OA publication miss part of the point. Most scholarship, including Near Eastern Archaeology, is manifestly not "financially sustainable." It depends on continued public and philanthropic support, and is not asked to turn a profit. It seems odd that scholarly communication, which is an important and essential aspect of the research process, should be asked to be profitable in disciplines dependent on public subsidies. If the expensive part of the equation (knowledge production) is already subsidized, why charge for the side that is nearly free (knowledge dissemination)? This is not a claim that *all* archaeological publication should be free. Certain venues for scholarly communication have much higher production values than others. Charging for such publications is more appropriate. However, many archaeological publications

rely on "professional volunteers" who write and edit for recognition, not direct financial rewards. They emphasize academic content over high value design, artwork, and layouts. These venues can readily adopt OA models, and in practice, a growing number of OA journals have found different ways to sustain themselves. Most of such journals are free for both authors and readers. For example, the nonprofit Scholarly Exchange offers an e-journal hosting service for \$750 per year. This level of cost is well within the range of many academic departments (and substantially less than some institutional journal subscription costs).

Moreover, OA can be achieved even without changing subscription protocols. The growing practice of "self-archiving" (uploading pre-prints or post-prints of papers to websites, as permitted by most academic publishers) is fueled because it offers similar advantages in increased citation as publishing in an OA journal (Harnad and Brody 2004; Hajjem, Harnad, and Gingras 2005). Thus, OA is mostly likely to be achieved in Near Eastern archaeology first through the enlightened self-interest of authors. The increasingly popular practice of uploading journal articles to websites may well be the route to which OA becomes a common and expected part of our discipline.

An Open Future for the Past

OA models are proliferating, not only for sharing traditional forms of scholarly production (peer-review papers), but also new forms of content, especially databases and media archives. Open Context, though still in active software development, already has a variety of demonstration datasets now available for exploration and testing. Our primary goal now is to build the "critical mass" of users, contributors, and content needed to sustain Open Context as a valued scholarly resource. To this end, we are seeking from the Near Eastern archaeological community and beyond, excavation and survey data, media, museum and comparative collections and content reviewers. We are also seeking collaborations with other open-source projects to build links to other systems, including multi-language interfaces; regional, temporal, or subject-specific community portals; and additional data-versioning tools.

The future of the past is looking increasingly open. OA makes research easier to find and use, and gives that research more impact and significance. With sufficient community contributions, feedback, and support, Open Context and related OA systems can open new doors of understanding and facilitate collaborations across disciplines. Reference materials will be easier to find and use, excavations results will be easier to interpret and reanalyze, and broad regional syntheses will be more comprehensive and analytically rigorous. Across the board, OA now has a great deal of momentum and

powerful institutional support. We call on the Near Eastern archaeological community to participate in this fundamental transformation of scholarly practice.

Acknowledgments

The authors would like to thank those involved in the creation of Open Context, in particular David Schloen and the University of Chicago OCHRE project, for their continued support and partnership. Their efforts have provided the essential conceptual groundwork needed for developing the Open Context system. Special thanks also go to Benjamin Porter and Ahrash Bissell for their insightful comments on this paper. Needless to say, any omissions or errors in this work are solely the fault and responsibility of the authors. Finally, Open Context and many other open education initiatives directly result from the enabling financial support of the William and Flora Hewlett Foundation. Their support, together with the generous financial contributions of Doris and Donald Fisher and the Joukowsky Family Foundation, help ensure that Open Context can serve as a free and open access resource for the community.

Full Disclosure Matters

Jane M. Cahill and James A. Passamano

In the preceding article in this Forum, Sarah Whitcher Kansa, Eric C. Kansa, and Jason M. Schultz have announced that “a recently released data-sharing web application, called Open Context (www.opencontext.org), enables researchers to publish structured data along with textual narratives and media (images, maps, drawings, videos) on the web.” The authors explain that Open Context is designed to facilitate open-access publishing of archaeological data by capitalizing on advantages of “‘born digital’ documentation,” now commonly used by excavators to record vast amounts of data. Their article ends with an appeal for help to achieve their “primary goal . . . to build the ‘critical mass’ of users, contributors, and content needed to sustain Open Context as a valued scholarly resource.” Apparently for this reason, their description of Open Context is neither critical nor objective but, instead, aimed at persuading excavators to use Open Context to publish field data. The authors could have presented a more useful portrait of Open Context had they not focused exclusively on its perceived benefits but, instead, candidly addressed both the potential advantages and disadvantages that it poses for data providers and users.

Open access publishing holds great promise for the scholarly community in general and the community of Near Eastern archaeologists in particular. However, failing to realize that Open Context is more likely to be embraced by scholars who have been fully informed about both the benefits and risks of publishing data on the internet, Kansa, Kansa, and Schultz acknowledge that “technologies for . . . internet publication of archaeological data” pose “important social, legal, and incentive issues,” but fail to discuss those issues. Instead, they present a brief discussion of copyright to which the following points should be added: (1) copyright is a complex body of law that has become increasingly difficult to apply to new and emerging technologies like Open Context; and (2) Open Context’s copyright management measures are some-rights-reserved licensing agreements developed and distributed by Creative Commons whose Science Commons Project has recently released a Protocol for Implementing Open Access Data (Protocol) that recommends such licenses not be used for publishing data. (See <http://sciencecommons.org/projects/publishing/open-access-data-protocol/>; cited March 15, 2008.)

Archaeologists who read Kansa, Kansa, and Schultz’s description of copyright will likely be pleased to learn that Open Context allows them to retain copyright, but may incorrectly conclude that data published on Open Context will be entitled to copyright. Copyright is not likely to protect most archaeological field data because in the United States copyright extends only to original, that is, creative expressions, and does not extend to facts or ideas. The Kansas and Schultz’s conclusion to the contrary derives from their assumption that unprotected facts and ideas included in archaeological databases will be accompanied by expressive content such as images that are entitled to copyright. Nevertheless, data published on Open Context consisting of facts and ideas will belong to the public domain and be freely usable by others to create their own works including integrated databases and excavation reports, possibly even before the excavator who originally collected the data, found the facts, or formulated the ideas. Since the expectation of first publication provides substantial incentive for archaeological research, that incentive could be diminished—if not eliminated—for excavators who publish data from their excavations on Open Context before they have synthesized that data into an interpretive report.

Archaeologists who read Kansa, Kansa, and Schultz’s description of copyright may also conclude that Open Context’s copyright management measures will not only protect their copyright interests but also promote open access to archaeological data. They state that

Open Context, like many other open access (OA) systems, has adopted policies requiring use of standard copyright licenses that give explicit permissions for reproducing content. The non-profit

and increasing capacities—broadens the potential for showing large numbers of finds. For example, a digital reporting (<http://www.cobb.msstate.edu/dignew/start.htm>) of ceramic and stone figurines found at Tell Halif in Israel included detailed photographic and analytical data, including 307 virtual movies, of all 850 fragments. Should the conclusions about the significance of this collection of figurines prove to be questionable—as it surely will as more evidence comes to light—future scholars will have the benefit of a full disclosure of the data.

Furthermore, because some items of this collection have already migrated from the archives of the Israel Antiquities Authority to museum displays, and since the entire collection has already been given catalog identifications in a system different from that applied in the field, it seems likely that future scholars may have to depend in some instances (individual pieces will disappear) entirely on the digital record alone.

Collaborating with Future Archaeology

Because present conclusions become the building stones of the future, because assured results and hard-won consensus will be challenged and altered, excavators need to perceive a significant element of their work as preparing accurate and plentiful data that will be used in new and exciting conclusions otherwise not imagined by themselves. We ought to be aware of our work as something more than publications that count toward tenure, but also as something that will serve future research. If this claim is true, then how and what we report, how we record data, in fact, even how and what we collect will be impacted by the end-game of publishing data for collaborative work with future research. Because digital distribution and digital publication has made it feasible, an excavator may now show in graphic format more than the selected representative examples of items that support the conclusion she has reached; indeed, it may be more important for the collaboration with future research that those items that played little or no role in support of the conclusion also be published. That is, it will be equally as important to share the raw data as it is to present the items selected to demonstrate the case. All of the data needed for the success of future excavation must be liberated from the archives of antiquities authorities and from the personal libraries of original excavators.

This capacity to publish (digitally) fuller sets of data allows scholars to reject traditional conditions of passivity in receiving pre-selected and pre-classified information; scholarship will come to expect access to original, unselected data and will organize them in alternate and novel ways.

Reference

- Anderson, C.
2006 *The Long Tail: Why the Future of Business is Selling Less of More*. New York: Hyperion.
- Bearman, D., and Trant, J.
2005 Social Terminology Enhancement through Vernacular Engagement: Exploring Collaborative Annotation to Encourage Interaction with Museum Collections. *D-Lib Magazine* 11(9). Online: <http://www.dlib.org/dlib/september05/bearman/09bearman.html>.
- Benkler, Y.
2006 *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press. Online: http://www.benkler.org/Benkler_Wealth_Of_Networks.pdf.
- Brown, G. O.
2003 Out of the Way: How the Next Copyright Revolution Can Help the Next Scientific Revolution. *PLoS Biology* 1(1):e9. Online: <http://biology.plosjournals.org/perlserv?request=get-document&doi=10.1371/journal.pbio.0000009>.
- Denning, K.
2004 "The Storm of Progress" and Archaeology for an Online Public. *Internet Archaeology* 15. Online: http://intarch.ac.uk/journal/issue15/denning_index.html.
- Giles, J.
2005 Internet Encyclopaedias Go Head to Head. *Nature* 438 (issue 7070): 900–901. Online: <http://www.nature.com/news/2005/051212/full/438900a.html>.
- Harnad, S., and Brody, T.
2004 Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine* 10(6). Online: <http://dlib.org/dlib/june04/harnad/06harnad.html>.
- Hajjem, C.; Harnad, S.; and Gingras, Y.
2005 Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Engineering Bulletin* 28(4): 39–47. Online: <http://eprints.ecs.soton.ac.uk/11688/>.
- Hodder, I.
1999 Archaeology and Global Information Systems: *Internet Archaeology* 6. Online: http://intarch.ac.uk/journal/issue6/hodder_toc.html.
- Jacobs, P.
2001 Ancient World, Digital World: Excavation at Halif. *Ariadne* 27 (March). Online: <http://www.ariadne.ac.uk/issue27/jacobs/>.
- Jacobs, P., and Holland, C.
1997 Archaeology Online: New Life for Old Dead Things. *First Monday: Peer Reviewed Journal on the Internet* 2(6). Online: http://www.firstmonday.dk/issues/issue2_6/jacobs/index.html.
- Jacobs, P., and Tsioukas, V.
2004 Toward a Digital Collection: The Lois Dowdle Cobb Museum. *Museology* 1. Online: <http://www.aegean.gr/culturaltec/museology>.

- Kansa, E. C.
2005 A Community Approach to Data Integration: Authorship and Building Meaningful Links across Diverse Archaeological Data Sets. *Geosphere* 1(2):97–109. Online: <http://geosphere.geoscienceworld.org/cgi/content/full/1/2/97/>.
- Kansa, E. C.; Schultz, J.; and Bissell, A.
2005 Protecting Traditional Knowledge and Expanding Access to Scientific Data. *International Journal of Cultural Property* 12(3):285–314.
- Kintigh, K.
2006 The Promise and Challenge of Archaeological Data Integration. *American Antiquity* 71(3):567–78.
- National Science Foundation
2007 Data, Data Analysis, and Visualization (2006 – 2010). In *Cyberinfrastructure Vision for 21st Century Discovery* (NSF 07-28, March 2007). Washington, DC: National Science Foundation. Online: <http://www.nsf.gov/pubs/2007/nsf0728/index.jsp>.
- Piwowar, H. A.; Roger S. D.; and Fridsma, D. B.
2007 Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS One* 2(3):e308.
- Ravindranathan, U.; Shen, U. R.; Gonçalves, M. A.; Fan, W.; Fox, E. A.; and Flanagan, J. W.
2004 ETANA-DL: a digital library for integrated handling of heterogeneous archaeological data. *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries: Global Reach and Diverse Impact, Tucson, Arizona, June 07–11, 2004*. New York: ACM Press. Online: <http://portal.acm.org/citation.cfm?doid=996350.996370>.
- Schloen, J. D.
2001 Archaeological Data Models and Web Publication Using XML. *Computers and the Humanities* 35:123–52.
- Society for American Archaeology
1996 Principles of Archaeological Ethics. *Society for American Archaeology*. Updated January 18, 2007. Online: <http://www.saa.org/ABOUTSAA/COMMITTEES/ethics/principles.html>
- Trant, J.
2006 Exploring the Potential for Social Tagging and Folksonomy in Art Museums: Proof of Concept. *New Review of Hypermedia and Multimedia* (in press). Online: <http://www.archimuse.com/papers/steve-nrhm-0605preprint.pdf>.
- Weinberger, D.
2007 *Everything is Miscellaneous: The Power of the New Digital Disorder*. New York: Holt.
- Willinsky, J.
2006 *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: MIT Press. Online: <http://mitpress.mit.edu/catalog/item/default.asp?tttype=2&tid=10611>.

American Schools of Oriental Research

2008 ANNUAL MEETING



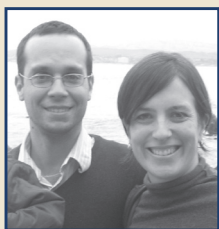
November 19–22
Westin Waterfront Hotel

BOSTON MASSACHUSETTS

Supersaver registration
deadline is Nov. 3, 2008

For more information:
<http://www.asor.org/AM/am.html>

ABOUT THE AUTHORS



Sarah Whitcher Kansa is Executive Director of the Alexandria Archive Institute (AAI), a San Francisco-based non-profit organization working to develop innovative digital services for openly sharing archaeological research. She is also a practicing zooarchaeologist whose long-term research has focused on

developments in social complexity and animal exploitation at the Late Neolithic site of Domuztepe, south-central Turkey.

Eric C. Kansa is Executive Director of the Information and Service Design (ISD) program at UC Berkeley's School of Information. The ISD Program supports research in "service science" and a clinical program that integrates research and instruction in collaborative IT projects. He is currently Convener of the Society for American Archaeology's Digital Data Interest Group and is a member of the Disciplinary Advisory Board for Archaeoinformatics.org.



Jason M. Schultz is Associate Director of the Samuelson Law, Technology & Public Policy Clinic and a faculty member at the U.C. Berkeley School of Law. Previously he was a senior staff attorney at the Electronic Frontier Foundation, a leading digital rights group, where he handled high-profile

intellectual property and technology matters affecting the public's interests in free expression, fair use, and innovation. His writing focuses primarily on the ongoing struggle to balance intellectual property regimes with the public's interest in free expression, access to knowledge, and innovation in light of new technologies and the challenges they pose. He has been an advisor on these issues for the Open Context project since its inception.



Jane M. Cahill has a B.A. and a J.D. from the University of Texas at Austin, and an M.A. from the Hebrew University where she is a Ph.D. candidate at the Institute of Archaeology. She has published numerous articles on a variety of topics, and is employed as a career

law clerk by United States District Court Judge Simeon T. Lake, III, in Houston, Texas.



James Passamano has a B.A. from Emory University, a J.D. from South Texas College of Law, and a LL.M. from Cambridge University. He is a principal in the Houston, Texas based law firm of Sufian & Passamano with seventeen years of practical experience

that includes representing a broad range of clients, publishing on a variety of legal subjects, and teaching as an adjunct professor at the University of Houston, Rice University, and South Texas College of Law.



Mitchell Allen is founder and publisher of Left Coast Press, Inc., a scholarly press specializing in archaeology, anthropology, and museum studies. He previously founded and directed AltaMira Press and has spent thirty years in academic publishing. Allen has a Ph.D. from UCLA, has

done fieldwork in Israel and Afghanistan, and currently teaches archaeology at Mills College in Oakland, CA.



Paul Jacobs teaches at Mississippi State University and is on the staff of the Cobb Institute of Archaeology. As a Senior Staff member of the Lahav Research Project, Jacobs has directed excavations at Tell Halif. With Chris Holland he has developed internet

resources for archaeology, which may be seen at <http://www.cobb.msstate.edu/Research.html>.



Christopher Holland has had over fourteen years experience in dealing with internet data applications. He has designed and built multi-million dollar corporate e-commerce systems and has headed research projects with Mississippi State University, the Smithsonian Institution, and the U.S.

Navy's COAST oceanographic education program. He is currently developing software for the Cobb Institute of Archaeology and the popular internet site LibraryThing.