Open Data on the World Wide Web: Opencontext.org and Data Integration as a Community Process

Eric C. Kansa
The Alexandria Archive Institute
San Francisco, CA (USA)
ekansa@alexandriaarcive.org

The Internet and Scholarly Communication

Open Access (OA) systems for scholarly communication have made impressive strides in recent years. Several studies now clearly document that self-archiving and open access publication enhances uptake and citation rates [Hajjem et. al. 2005]. Researchers therefore enhance their reputation and stature by opening up their scholarship. Mounting pressure for greater public access also comes from public backers of research. Granting foundations and government agencies interested in maximizing the return on their basic research investment often encourage and increasingly require some form of open access electronic dissemination. In the United States, the "Federal Research Public Access Act," would require OA for drafts of papers that pass peer review and result from federally funded research [Library of Congress 2006]. The bill would create government-funded digital repositories that would host and maintain these draft papers. The European Union is even further down the OA road, with the European Commission aggressively moving forward with the development of OA policies [Van Orsdel & Born 2007; Harnand 2007]. University libraries are some of the most vocal advocates for OA research. Current publishing frameworks have seen dramatically escalated costs, sometimes four times higher than the general rate of inflation [Createchange.org 2003]. Increasing costs have forced many libraries to cancel subscriptions and thereby hurt access and scholarship [ACRL 2003, Suber 2004]

Beyond Papers: Sharing Other Scholarly Media

OA models are gaining traction for sharing other types of scholarly content, beyond peer-reviewed papers. 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 research documentation in many areas of "small science". "Small Science", a term coined in this context by Paul Uhlir and Peter Schröder [cited by Onsrud and Campbell 2006], typically sees research conducted by small teams or individual investigators. More efficient and comprehensive sharing of the primary data collected in the context of small science is an important goal for many seeking to reform and enhance scholarly communication [ARL 2006: 57-59]. As demonstrated in ecology and other sciences, reused primary data can be an important resource for advances in understanding [Kansa 2005].

Data sharing does present a new set of technical, conceptual, and incentive problems. Small science typically works with very case-specific research questions, often using customized methodologies and recording systems. Archaeology is a good example, where field

documentation strategies are shaped by highly variable research agendas. For example, one excavates a Paleolithic cave site in a very different manner than a Roman urban site. Also, social and political factors external to scientific aims are very important in shaping documentation strategies. The majority of archaeological research takes place in the context of cultural resource management, where archaeological investigations are legally mandated to mitigate damage to historical and archaeological resources that may be caused by construction and other development. Excavation sampling strategies and the types of laboratory analyses conducted, may all be shaped by construction timelines and imperatives, permitting requirements, property owners, and community interest groups.

As a consequence, archaeological excavation results, specialist analyses and museum collection databases are highly variable [Kintigh 2006]. Recording and documentation strategies are typically customized for each individual study. Because archaeology depends on inputs from several different scientific disciplines (zoology, botany, geology and various environmental sciences, to name a few), archaeological datasets can be complex. At the same time, new archaeological research questions often demand more comprehensive and thorough documentation. Stratigraphic recording, intensive intrasite and regional survey techniques, photography (and video), geographic information systems (GIS), and computer aided design (CAD) all add to the richness and comprehensiveness of contemporary archaeological recording. While this information is never wholly complete or objective, it presumably has some value [Richards 2003].

This diversity and complexity presents an important challenge for meaningful data sharing. One cannot simply put data files on the Internet and expect them to be of immediate use to the community. Such datasets typically need extensive documentation to be intelligible to other researchers. Even *if* such metadata documentation is available, it is still difficult for a researcher to know if a given dataset is of interest. Even well documented datasets require several steps (downloading, launch of the proper software application, and user orientation) before they can be thoroughly investigated.

Maximizing the usability of primary research data involves more than access and metadata documentation. Ideally, such datasets should be available for casual inspection and analysis without requiring the user to download individual data files or launch special software. This level of access requires data integration strategies and a web-based infrastructure that can enable users to interact with pooled datasets. Fortunately, such systems are beginning to emerge in archaeology. The Etana Digital Library project (Etana-DL) < http://feathers.dlib.vt.edu:8080/etana/servlet/Start >, led by James Flanagan and digital library pioneer Edward Fox, has successfully demonstrated a data-mediation system that uses software to translate local data structures to a more general data structure. This mediation enables Etana-DL to provide interoperability and integrated search, browse, and analysis tools for several Near Eastern excavation datasets [Flanagan et al. 2004]. Dean Snow and colleagues advocate developing advanced text-mining systems to extract comparative data from archaeological reports, including “grey literature” documentation generated from cultural resource management (CRM) activities [Snow et al 2006]. Following the model of other scientific disciplines, the NSF recently awarded a group led by Keith Kintigh and colleagues a large “cyber-infrastructure” grant to stimulate data integration and sharing in archaeology. This project is now in its initial stages and aims to begin by developing ontologies for zooarchaeology. Ontologies are formally defined conceptual systems and are often used to support the integration of multiple datasets within a discipline. The CIDOC-CRM, a conceptual ontology important to many European cultural heritage (especially museum) initiatives, is perhaps the most significant ontology currently relevant to archaeology [Doer 2003].

An "*Open Context*" for Excavation Results and Related Collections

Other working systems are now coming online, including two related systems, the University of Chicago OCHRE project < http://ochre.lib.uchicago.edu/ > and *Open Context* < http://www.opencontext.org >. Both systems share the same data architecture described by the "Archaeological Markup Language" (ArchaeoML) and both have similar capabilities for integrating and pooling complex and media-rich archaeological documentation [Schloen 2001; Kansa 2005]. While OCHRE provides sophisticated data management tools targeted for active research projects, *Open Context* (Figure 1) is aimed at streamlined, web-based access and retrieval of field science and material collections-related content.

Figure 1. A simple map enables users to locate content based on geographic location

Open Context enables open public access to primary field data, notes, and media (images, maps, drawings, videos) on the web. It provides an easy to use, yet powerful, online database for exploring, searching, and analyzing multiple excavation results, survey datasets, and museum collections. Although initially intended to meet data dissemination needs of archaeology, we are

experimenting with using *Open Context* for other disciplines. As a result, *Open Context* now also offers public health and evolutionary biology datasets.

All of these diverse datasets can be explored by browsing through a map, or through different search options. *Open Context* is built with standard but powerful web technologies (MySQL and PHP), making it easy to integrate with a host of other web services, including weblogs, e-journals, and commercial search engines. These technologies help make *Open Context* accessible to search engine-indexing software. Search engine discovery will likely take a greater role in increasing the impact and uptake of research [Jensen 2005; Vaughan and Shaw 2005].

The Types of Material in *Open Context*

Open Context is best suited for publishing large bodies of complex structured data. All content is mapped to the ArchaeoML global schema, providing a common structure that integrates researcher datasets into a cohesive resource. This common structure simplifies the development of services around such diverse content. *Open Context*'s interface makes common browse, search, and analysis functions simple and straightforward. Users have a variety of options to find materials in *Open Context*, including simple, "Google-like" text searches to more sophisticated, advanced searches that use Boolean logic (Figures 2, 3 and 4). Simple charting tools help with data visualization (Figure 5), and a selected dataset (potentially drawing records from multiple projects) can be exported into common formats, such as MS-Excel.

Figure 2. A simple search for "carnelian"

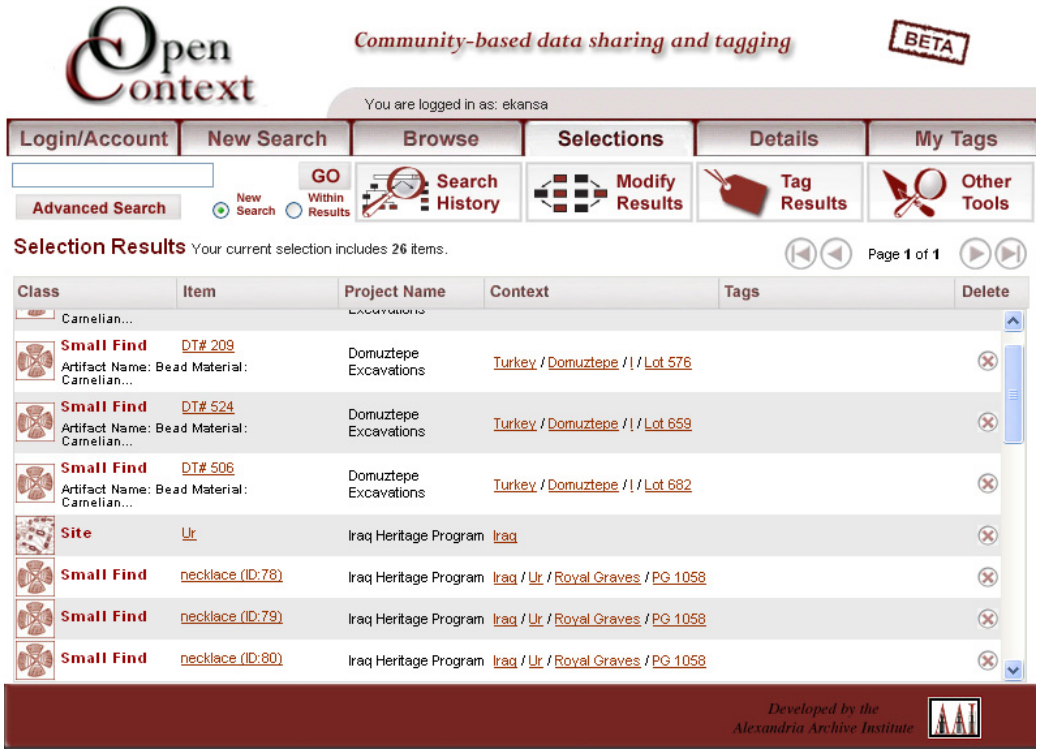

Figure 3: Results of the above "carnelian" search, showing items from multiple projects

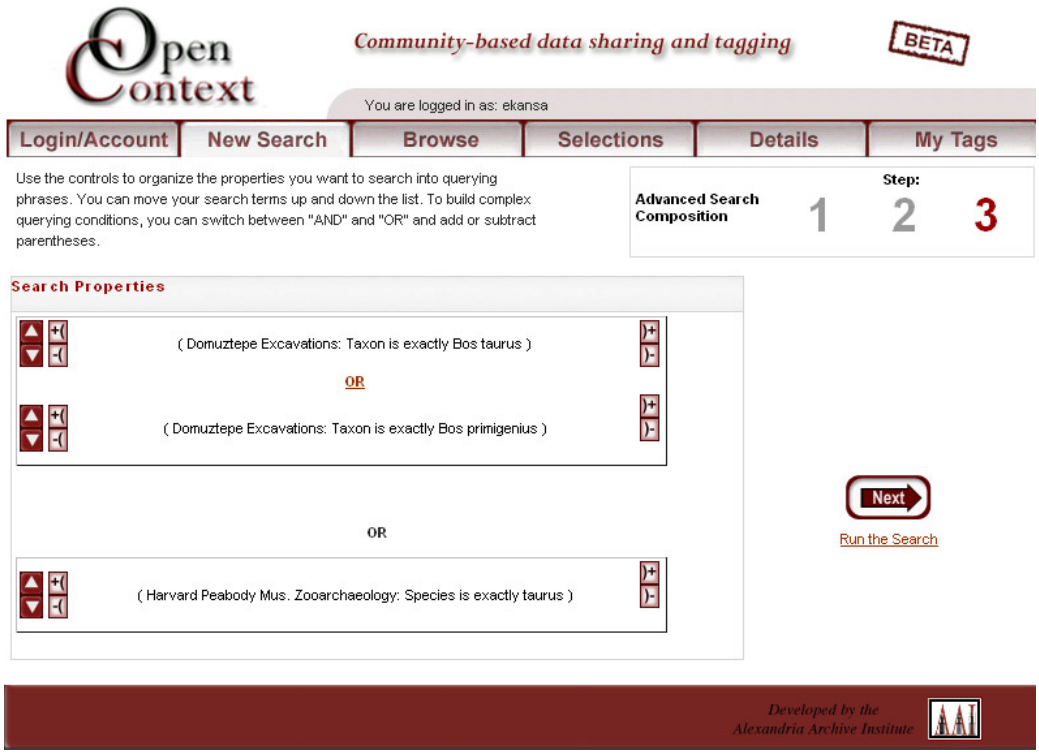

Figure 4. Example of an "Advanced Search" across multiple projects

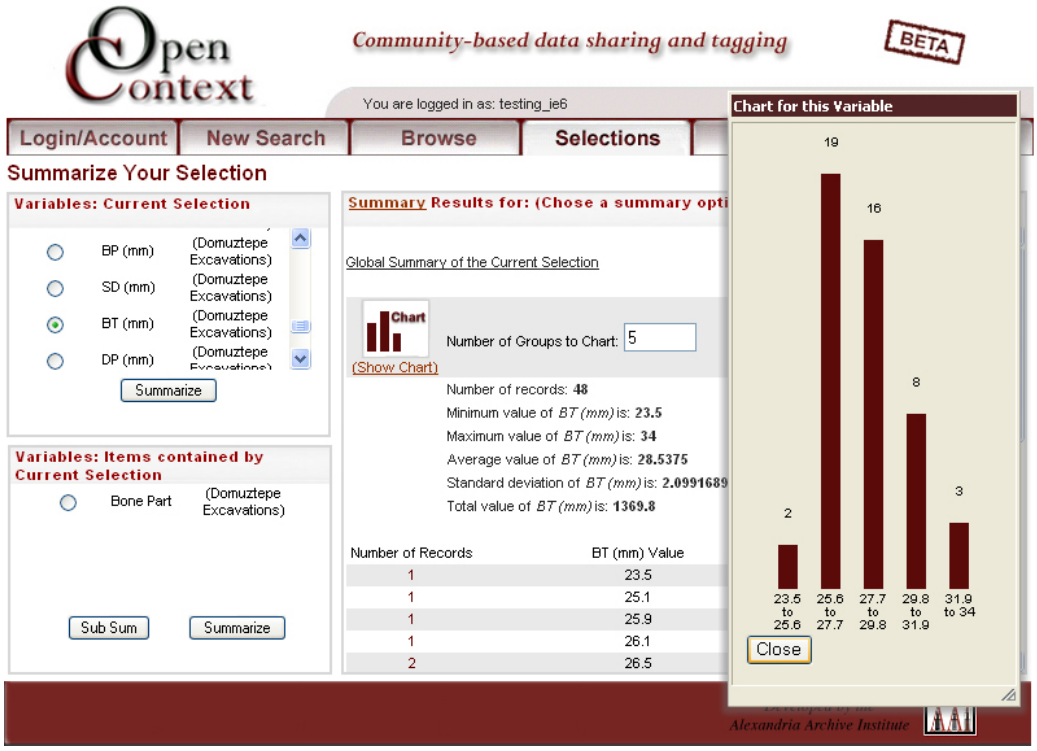

Figure 5. Example data visualization and charting

With these tools, users have open access to a variety of content. Archaeological research includes a wide variety of documentation, ranging from diaries and notes, media (photographs, drawings, video, GIS files, etc.) and more highly structured "database" content. Archaeological examples of structured data typically include: context databases, finds registries, museum registries and catalogues, and specialist analyses (Figure 6). Public health examples now include data on symptoms and disease causing agents of corneal ulcers. Evolutionary biology datasets include morphological measurements of different lizard populations. *Open Context* brings these diverse types of documentation together in a cohesive framework for easy community reuse and evaluation (Figures 7 and 8). Because *Open Context* was designed to meet the breadth and diversity of archaeological content, it can be adapted for other disciplines as well.

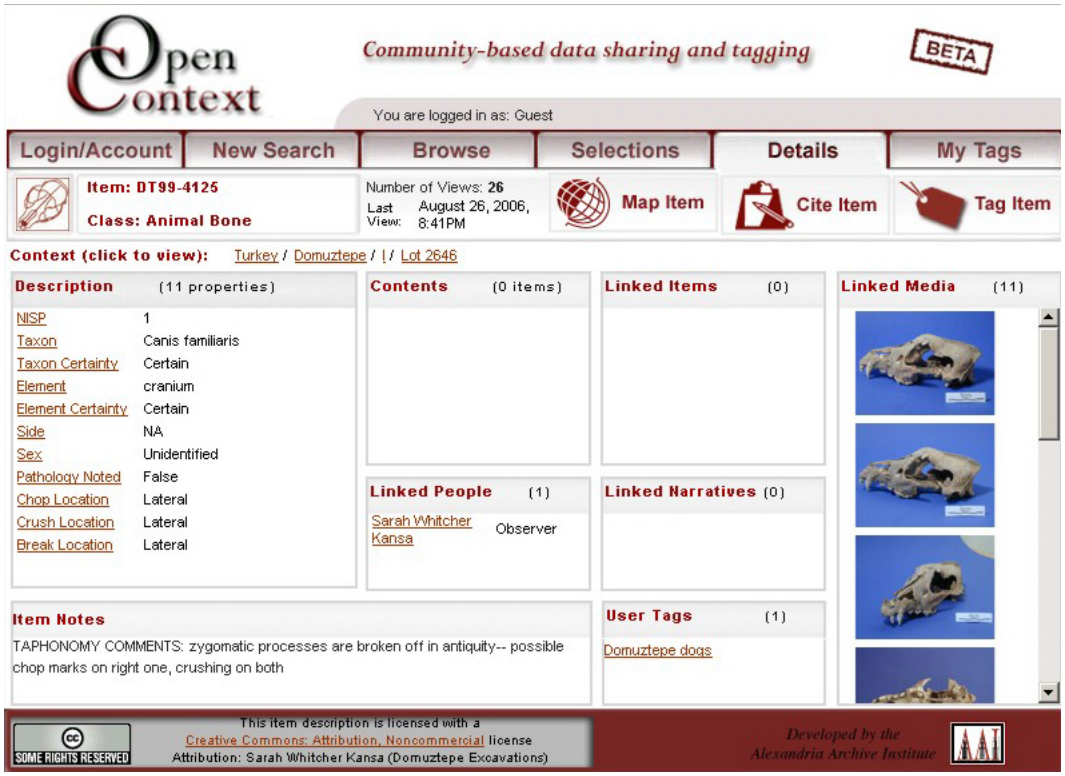

Figure 6. A record from the Domuztepe Zooarchaeological Analysis Database

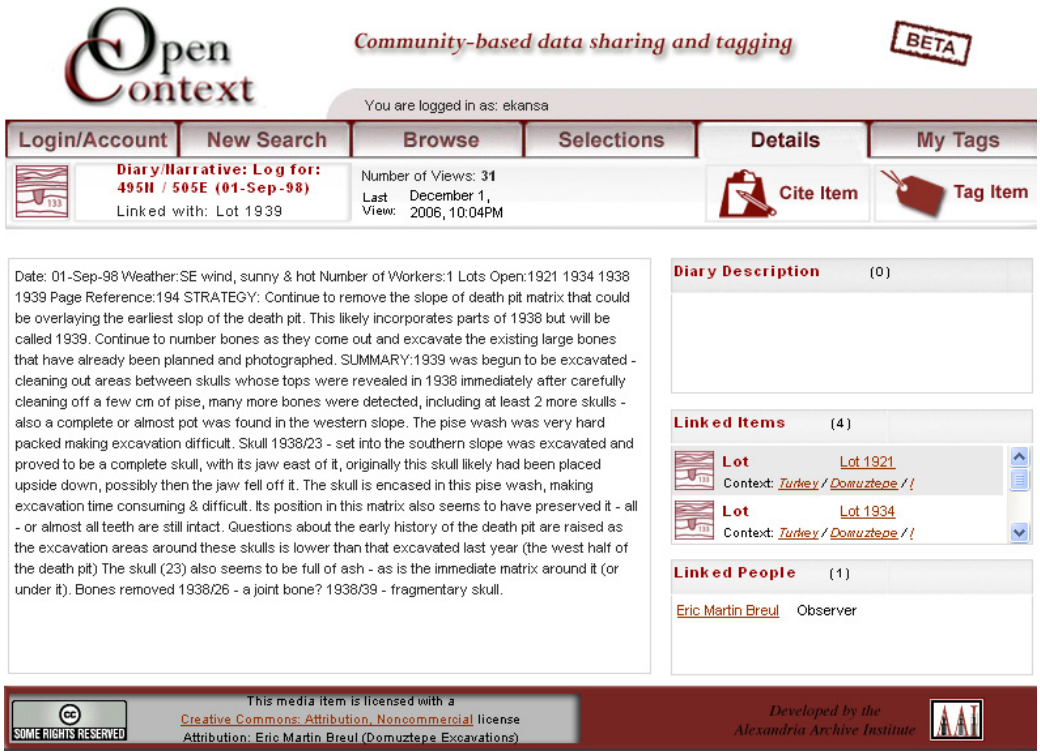

Figure 7. An excavation diary linking to items from the Domuztepe Context Database

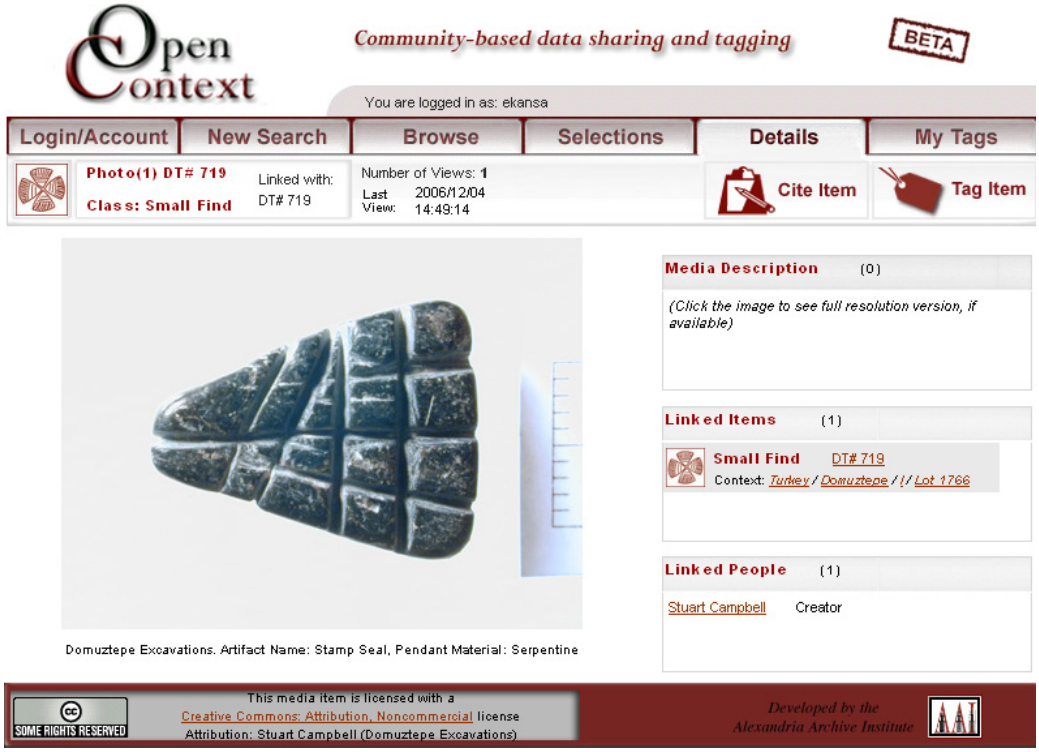

Figure 8. An image linked with its small finds registry record and context

Access, Copyright and Reuse of Content

Open Context is an open access publication system. All content is freely available on the World Wide Web. Because much of the content in *Open Context* has a high degree of originality in expression (in the sense of intellectual property law), copyright protections apply to much of *Open Context*'s content. To help keep this material open, *Open Context* follows a similar policy as the Public Library of Science (PLoS), requiring each contribution to carry an open, Creative Commons license. These licenses give explicit permissions for users to freely and legally use the material so long as they properly attribute the original creator [Brown 2003]. Creative Commons licenses include machine-readable RDF metadata that is captured by commercial search engines such as Yahoo and Google [Kansa et al. 2005]. This metadata facilitates discovery of openly licensed content, including *Open Context* resources. However, unlike PLoS, *Open Context* allows the full suite of Creative Commons licensing options, including use of the controversial "noncommercial" term. Many of the communities *Open Context* aims to serve have deep concerns over commercialization, especially with regard to antiquities. Therefore, despite the drawbacks and ambiguities of this licensing option, *Open Context* permits its use. Finally, *Open Context* has a policy that allows contributors to retain copyright to their content. This policy is intended to encourage dissemination through *Open Context* by not precluding publication in other more established venues (especially journals and books).

Creative Commons licenses allow great latitude in determining how "attribution" should take place. Since *Open Context* is intended to be a scholarly publication system, proper citation is the form of attribution required for reusing content published in *Open Context*. Clear citation will hopefully make researchers more comfortable with publishing in *Open Context*, since their contributions will be recognized in familiar ways. Clear citation also makes *Open Context* more

useful for scholarly applications. The system automatically generates citation information and a stable URL for each item in the database (Figure 8). Finally, the bibliographic metadata stored in *Open Context* is also expressed using the COinS ("ContextObjects in Spans", see < http://ocoins.info/ >) standard. COinS is a micro-format for expressing Dublin Core metadata and is readable by the new *Zotero* (< http://www.zotero.org >) citation tool. Using *Zotero*, investigators can automatically capture bibliographic information associated with *Open Context* materials.

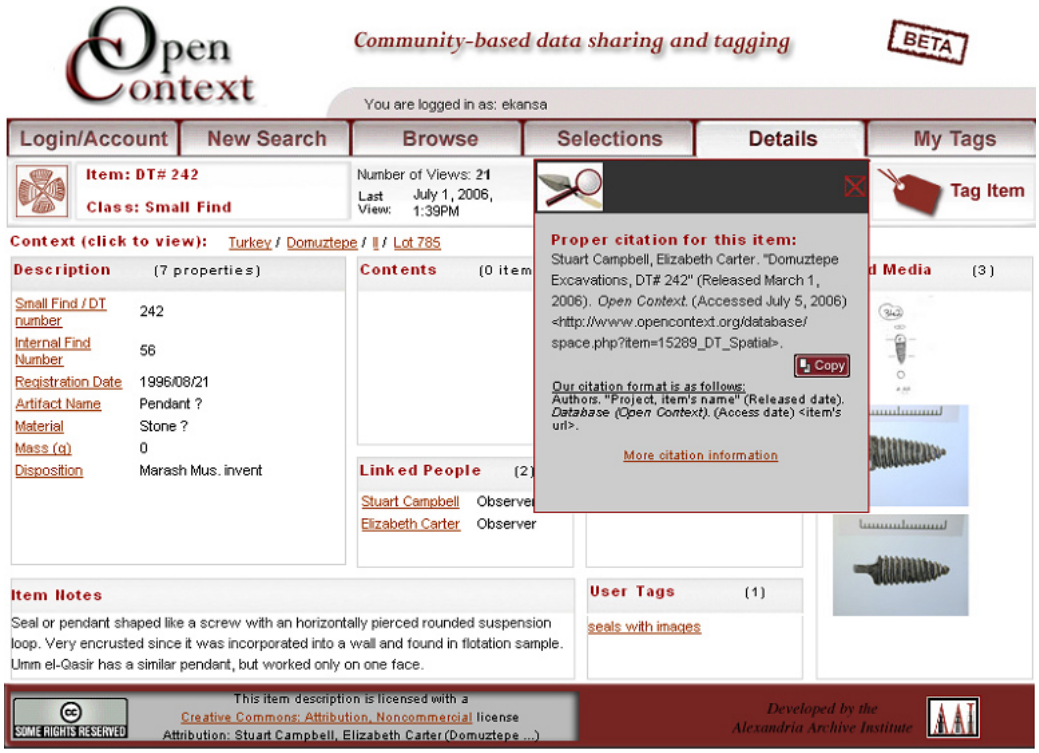

Figure 9. Automatically generated citation for an Open Context item

While *Open Context* offers several citation tools, web publishing of primary datasets raises some interesting questions about value and credit. An excavation dataset or raw data on animal population morphological variation are very different forms of scholarly contribution than a peer-reviewed paper. Refereed papers are currently the main (and sometimes only) currency of professional achievement. To begin assessing the scholarly impact of digital datasets, *Open Context* records information on visits to each record in the system. Recent studies have shown a significant correlation between download counts and more commonly used measures of citation impact in scholarly papers [Brody et al. 2006]. Nevertheless, a database in itself is a poor guide to understanding excavation or survey results. In order to be better understood and used, datasets are best linked with papers and narratives that synthesize observations and interpretations in a more meaningful framework [Richards 2003]. Similarly, the ability to reference public databases such as *Open Context* can enhance journal publications by making primary evidence more transparent and open for critical evaluation. Citation tools are help in this regard. However, technical and editorial coordination between scientific journals and online data repositories will do more to facilitate research. Finally, recent evidence suggests that peer-review papers containing primary data (and not only summaries of these data) have higher impact and citation rates [Piwowar et al. 2007]. Achieving higher impact and citation rates

should help motivate researcher to publish primary results using systems like *Open Context* along with more traditional peer-review papers. Impact incentives may also motivate journal publishers to collaborate with data sharing systems like Open *Context*.

To facilitate coordination with narrative syntheses, *Open Context* automatically generates reciprocal hyperlinks with other web services that support the open "ping-back" standard. If a person using a weblog or publishing in a "ping-back"-enabled e-journal references an item or a set of items in *Open Context*, the *Open Context* system will be automatically informed about what items are being referenced. *Open Context* will then display links back to the weblog post referencing the *Open Context* database. If desired, an editorial board can be assembled to subject all links to editorial review. This will ensure that *Open Context* only registers references to trusted sources. Conversations with developers of open source e-journal systems (such as those developed by the Public Knowledge Project < http://pkp.sfu.ca/?q=ojs >) show that ping-pack support is likely to be forthcoming. As e-journal systems gain popularity in the social sciences and other small science disciplines, such features will help ensure that *Open Context* users will easily find scholarly uses and interpretations of *Open Context* content.

Making Sense of Multiple Project Datasets

The lack of many formal standards in archaeology and many other small science disciplines makes data integration a challenge. *Open Context* uses special software for database schema mapping and import of researcher datasets. The current public datasets available on *Open Context* were imported by the *Open Context* development team using "in house" importer software. We are currently completing development of a web-based data-publishing tool called *Penelope* (Figure 10). *Penelope* enables individual contributors to upload their own data tables and submit them for review and publication in *Open Context*. Making schema mapping intelligible to contributors is a major challenge. To help meet this challenge, *Penelope* makes extensive use of the Dojo-AJAX framework (< http://dojotoolkit.org/ >). This framework enables *Penelope* to give users immediate and dynamic feedback about their schema mapping. Such immediate feedback will help users adjust import parameters as needed. Finally, since the ArchaeoML global schema is relatively simple, we hope the *Penelope* publishing tool will be easy enough to be used by a wide community of researchers. ArchaeoML's relative simplicity and generality may some advantages over the CIDOC-CRM (an important ontology for "cultural heritage" content) [Doer 2003; Kansa 2005]. CIDOC describes some 120+ relationship types, many with very specific and highly nuanced meanings. In contrast, ArchaeoML describes five major and very general conceptual categories, making it easier to implement in a community with little conceptual modeling expertise [Kansa 2005].

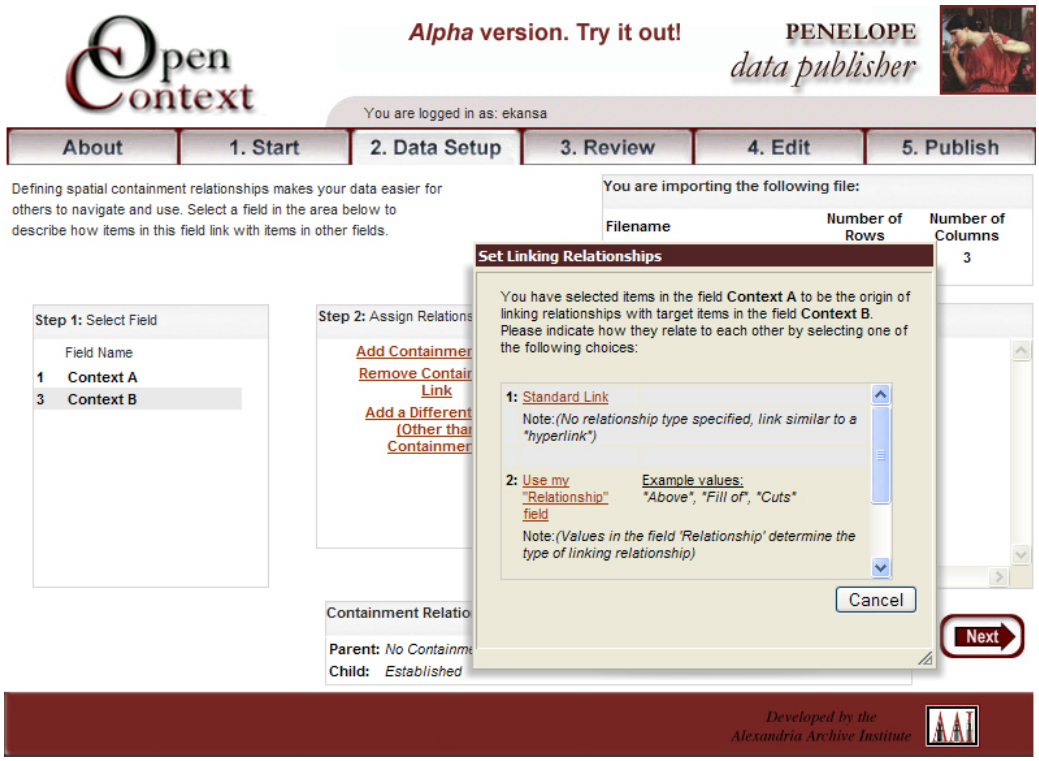

Figure 10. A user defining linking relationships using the Penelope data publishing tool

Penelope will make publishing with *Open Context* much more economical, since it will distribute the job of schema mapping across a wide community. Schema mapping and import of datasets through *Penelope* will also add significant value to datasets. *Penelope* asks users to describe their data, annotate them, indicate Creative Commons copyright licensing, and supply crucial metadata needed to facilitate community-wide reuse. Because *Open Context* uses ArchaeoML's very flexible architecture, all the original recording systems and terminologies and retained. Thus, *Penelope* enables researchers to publish their data in *Open Context* without forcing them to conform to overly restrictive predetermined standards. Moreover, *Penelope* stores schema-mapping parameters for each imported dataset. Essentially, it records how researchers think about their data and their data's organization. This information can be a valuable resource for developing conceptual ontologies, and may facilitate future automated text-mining and data-mining technologies.

To help make sense of this widely varying body of material, *Open Context* has a user "folksonomy" system. 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.

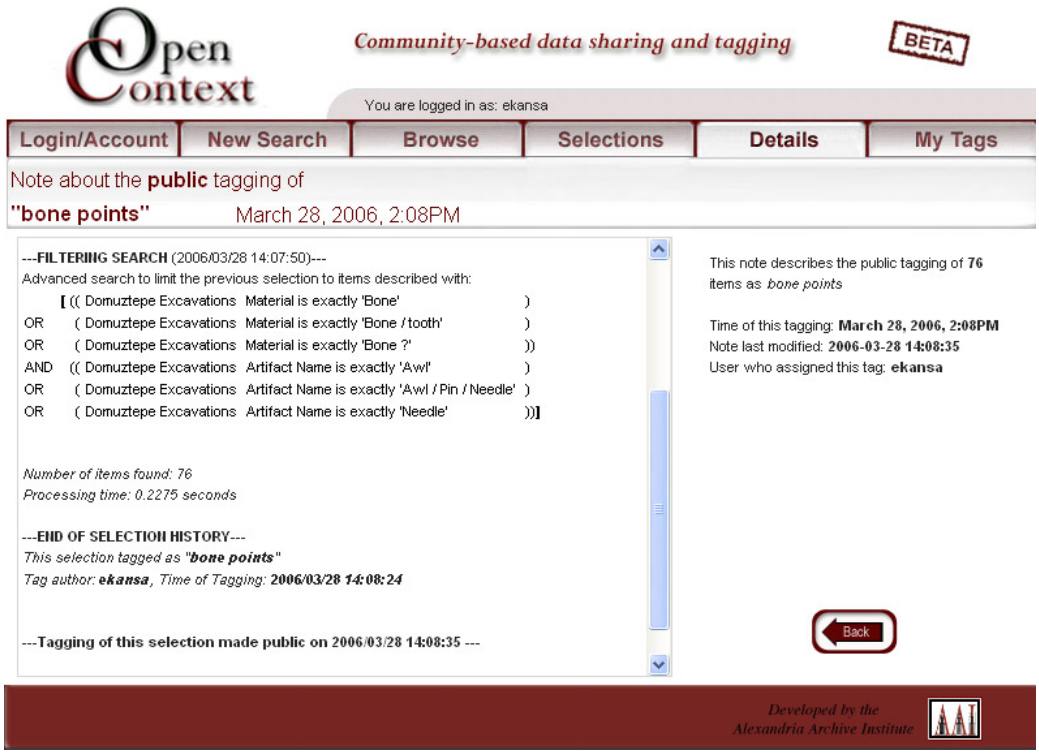

Figure 11. A search history saved with a tagged selection set

This folksonomy system can facilitate semantic data integration, and recent experiments suggest such systems offer annotations of sufficient quality to meet some needs of museum professionals [Bearman and Trant 2005; Trant 2006]. *Open Context* is developing several enhancements to this system, including better ways of recognizing professional credentials and scholarly. 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 shortly be enhanced to better recognize scholarly credentials and authority. We also anticipate proving options for users to apply professionally developed standard vocabularies such as the Getty Art and Architecture Thesaurus or future conceptual ontologies [see Kintigh 2006], including the CIDOC-CRM.

Looking Forward

The obvious advantage of *Open Context* is that a common database tool can disseminate many project datasets. This has considerable cost advantages. Researcher data can now be accessed, browsed, queried, and analyzed dynamically without building custom web databases for each project. Moreover, users can work with multiple datasets using a common set of tools and interfaces. *Open Context*'s flexible data structure can enable researchers to work across disciplinary boundaries and explore linkages between archaeological, environmental, historical, sociological and other types of structured information. Relating such varied forms of documentation may open doors for innovative research agendas. To help demonstrate and refine the applicability of *Open Context* to support cross-disciplinary research, its developers invite contributions of data and media from other small science fields needing better data dissemination.

Open Context must now build a "critical mass" of users, contributors and content needed to sustain *Open Context* as a valued scholarly resource. As *Open Context*'s collections and usage grow, its base of technical, administrative, and editorial support must also go. *Open Context* needs editorial assistance to help insure quality and oversee revision and error correction. Additional support may be provided through open source software development partnerships. Finally, data sharing and collaboration between multiple repositories is an important digital data longevity strategy [Reich and Rosenthal 2001]. Currently, the OCHRE project provides digital longevity support for *Open Context* pilot projects, and additional archival partnerships will be needed to better secure the often-irreplaceable archaeological content hosted by *Open Context*. To learn more about contributing to *Open Context*, please contact the author via email at (ekansa@alexandriaarchive.org).

With sufficient community contributions, feedback, and support, *Open Context* and related open access systems will expedite and streamline reference searches and provide a comparative format for efficiently interpreting and reanalyzing research results. Similarly, making primary data efficiently accessible and usable can support research agendas that are not currently achievable. By pooling primary data resources in systems with powerful analytic tools, such systems should enable broad regional syntheses that are more comprehensive and more analytically rigorous than are currently feasible [Kansa 2005; Kintigh 2006].

Acknowledgements:
The author would like to thank everyone who has participated in making Open Context operational, especially 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 Open Context development. Thanks must also be given for the collaboration of John Wilbanks (Science Commons), and Heather Ford (iCommons) for facilitating Open Context and other open science and open culture initiatives. Sarah Whitcher Kansa deserves special thanks for her comments, ideas, and edits. Needless to say, any omissions or errors in this work are solely the fault and responsibility of the author. 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, as well as, 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.

References Cited

Alliance for Tax Payer Access
2006 Federal Research Public Access Act (S.2695). Taxpayeraccess.org. Alliance for Tax Payer Access. < http://www.taxpayeraccess.org/frpaa/ >

Association of College and Research Libraries (ACRL)
2003 Scholarly Communication Toolkit: Journal Pricing. Ala.org. Association of College and Research Libraries (ACLR).
<http://www.ala.org/ala/acrl/acrlissues/scholarlycomm/scholarlycommunicationtoolkit/administrators1/admineconomics.htm >

Association of Research Libraries (ARL)
2006 To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering. ARL.org / Public Policies. Association of Research Libraries (ARL) < http://www.arl.org/bm~doc/digdatarpt.pdf>

Bearman, David, and Jennifer Trant
2005 Social Terminology Enhancement through Vernacular Engagement: Exploring Collaborative Annotation to Encourage Interaction with Museum Collections. *D-Lib Magazine* 11(9) <http://www.dlib.org/dlib/september05/bearman/09bearman.html >

Brody, Tim, Stevan Harnad, and Les Carr
2006 Earlier Web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology* 57(8):1060-1072.

Brown, Glenn Otis.
2003 Out of the way: How the next copyright revolution can help the next scientific revolution. *PLoS Biology* 1(1):e9 < http://biology.plosjournals.org/perlserv?request=get-document&doi=10.1371/journal.pbio.0000009 >

Createchange.org
2003 Scholars Under Siege: The Scholarly Communication Crisis. Createchange.org. Association of Research Libraries (ARL), Association of College and Research Libraries (ACRL), and the Scholarly Publishing and Academic Resources Coalition (SPARC) <http://www.createchange.org/archive/faculty/issues/silent.html>

Doerr, M.,
2003 The CIDOC conceptual reference module: An ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3): 75–92.

James W. Flanagan, Edward A. Fox, Douglas R. Clark, Weiguo Fan, Ravindranathan, Rao Shen, and Marcos André Gonçalves
2004 "Jordanian archaeology and ETANA: Developing a digital library for Near Eastern archaeology" in *History and archaeology of Jordan: Petra, Jordan*, ETANA Publications, p. 1–16, <http://feathers.dlib.vt.edu/publications.html>

Harnand, Stevan,
2007 The European Commission has already mandated Green OA! - Open Access Archivangelism, *Open Access Archivangelism*, <http://openaccess.eprints.org/index.php?/archives/223-The-European-Commission-has-already-mandated-Green-OA!.html > (accessed April 18, 2007).

Harnad, Stevan., and Tim Brody
2004 Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine 10(6)* < http://dlib.org/dlib/june04/harnad/06harnad.html>

Hajjem, Chawki., Stevan Harnad, and Yves Gingras
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) pp. 39-47. <http://eprints.ecs.soton.ac.uk/11688/>

Jensen, Michael
2005 Presses Have Little to Fear from Google, *The Chronicle for Higher Education* 51(44):B16 <http://chronicle.com/prm/weekly/v51/i44/44b01601.htm>

Kansa, Eric
2005 A community approach to data integration: Authorship and building meaningful links across diverse archaeological data sets. *Geosphere* 1(2):97–109

Kansa, Eric.C., Jason Schultz, and Ahrash Bissell

2005 Protecting Traditional Knowledge and Expanding Access to Scientific Data. *International Journal of Cultural Property* 12(3):97–109

Kintigh, Keith
2006 The Promise and Challenge of Archaeological Data Integration. *American Antiquity* 71(3):567–578

Library of Congress
2006 Federal Research Public Access Act of 2006 (Introduced in Senate). Library of Congress. < http://thomas.loc.gov/cgi-bin/query/z?c109:S.2695:>

Onsrud, Harlan and James Campbell
2006 Big Opportunities in Access to "Small Science" Data. *Commons of Science Conference Vision Paper*, < http://www.spatial.maine.edu/icfs/Onsrud-Campbell%20Paper.pdf> (accessed March 11, 2007).

Piwowar, Heather A., Roger S. Day, and Douglas B. Fridsma,
2007 Sharing Detailed Research Data Is Associated with Increased Citation Rate, *PLoS One* 2(3): e308, http://www.plosone.org/article/fetchArticle.action?articleURI=info:doi/10.1371/journal.pone.0000308 (accessed April 20, 2007).

Public Library of Science (PLoS)
2006 About *PLoS Biology*. Journals.plos.org. Public Library of Science. <http://journals.plos.org/plosbiology/information.php>

Richards, Julian
2003 Online archives. *Internet Archaeology* 15 <http://intarch.ac.uk/journal/issue15/richards_toc.html>

Reich, Vicky, and David S. Rosenthal
2001 LOCKSS: A Permanent Web Publishing and Access System. *D-Lib Magazine* 7(6) <http://www.dlib.org/dlib/june01/reich/06reich.html>

Science Commons
2006 Scholar's Copyright Project: Background. Sciencecommons.org. Creative Commons and Science Commons. <http://sciencecommons.org/projects/publishing/background.html>

Schloen, David
2001 Archaeological data models and Web publication using XML. *Computers and the Humanities* 35:123–152

Dean R. Snow, Mark Gahegan, C. Lee Giles, Kenneth G. Hirth, George R. Milner, Prasenjit Mitra, James Z. Wang
2006 Cybertools and Archaeology. *Science* 311(5763): 958 – 959.

Suber, Peter
2004 SPARC Open Access Newsletter, (April 2, 2004: Issue#72) <http://www.earlham.edu/~peters/fos/newsletter/04-02-04.htm>

Trant, Jennifer
2006 Exploring the potential for social tagging and folksonomy in art museums: proof of concept. New Review of Hypermedia and Multimedia 12(1):83-105

<Full-text preprint: http://www.archimuse.com/papers/steve-nrhm-0605preprint.pdf>

Van Orsdel, Lee C and Kathleen Born
2007 Serial Wars - 4/15/2007 - *Library Journal, LibraryJournal.com*, http://www.libraryjournal.com/article/CA6431958.html (accessed April 18, 2007).

Vaughan, Liwen, and Debora Shaw
2005 Web citation data for impact assessment: A comparison of four science disciplines. *Journal of the American Society for Information Science and Technology* 56(10):1075-1087