WHITE PAPER

APPLYING LINKED OPEN DATA: REFINING A MODEL OF DATA SHARING AS PUBLICATION

A Digital Humanities Implementation grant

Grant #: HK-50037-12

Project Director: Eric C. Kansa

Grantee Institution: The Alexandria Archive Institute

Submitted on December 30, 2015

to the NEH Division of Digital Humanities

PROJECT GOALS

The Alexandria Archive Institute received a Digital Humanities Implementation Grant of $261,056 for the project *Applying Linked Open Data: Refining a Model of Data Sharing as Publication* (Grant #: HK-50037-12). The goal of the project was to develop a model of "data sharing as publication" as a means to preserve quality archaeological datasets that are at risk of loss. This approach sees the publication of editorially-processed and peer-reviewed datasets as key to making data sharing and preservation an accepted norm.

Over three years (Sept. 2012 – Aug. 2015), this project used case studies to demonstrate an extensible model for publishing well-documented and reusable scholarly data in the Web-based, open-access data publishing system Open Context (http://opencontext.org; Figures 1-2). Major outcomes of the project, which are described in this report, include the following:

- Refinement of Open Context's capabilities to enable researchers to relate datasets with other data published on the Web using Linked Open Data methods (see Figures 3-5).
- Publication of high-quality datasets.
- Development of tools and workflows for researchers to relate datasets with other data published on the Web (using APIs and Linked Open Data methods).
- Alignment of data to common standards (ontologies) to facilitate research across multiple dataset.
- Improved understanding of how to manage data publication workflows and account for their costs by publishing datasets and observing the collaborative work of the researchers.
- Generation of a large body of data, publications, and data integration methods that will capture wide interest among archaeologists working in the Near East, East Mediterranean, and beyond.

The outward orientation of this project is a key innovation in perspective. Rather than aiming for centralization, this approach enables participation in a growing and widely-distributed humanities information ecosystem. Linked Open Data offers a powerful means to leverage distributed data for research applications. Innovative publishing workflows will contribute to the Web of Data, give researchers recognition and rewards, and open new research frontiers. This project contributes to the shared infrastructure of the humanities by demonstrating how publication processes can improve the discoverability, reuse, and longevity of primary scholarly materials. The datasets created by individual researchers still largely remain hidden, underutilized, and vulnerable to loss. Researchers in many natural science, social science and humanistic domains need workflows and venues to publish high-quality data. Venues that offer professional recognition and help to transform datasets into more powerful Linked Open Data will open new opportunities for scholarship. Thus, this project's outcomes are relevant to any field needing better data dissemination.

PROJECT ACTIVITIES AND ACCOMPLISHMENTS

OVERVIEW

The Implementation Grant activities described here took place over three years. The first phase (approximately six months) involved establishing the project's working groups and laying the groundwork for the publication and integration of related datasets. Activities involved soliciting working group participants, identifying datasets for publication, discussing data publishing processes, and exploring Linked Open Data options for the project datasets. For those datasets already finalized and ready for dissemination, we edited, reviewed, and annotated them and published them in Open Context. During this phase, we also gathered feedback from participating data authors on their needs around data publishing, specifically regarding the facility of publishing their own datasets, how to determine the quality of other researchers' data, and obstacles encountered when attempting to reuse and integrate multiple datasets. Technology developments during this period included a new feature in Open Context for previewing publication data to provide data contributors with a "page proof" version of their dataset before it is formally published.

During the next six months of the project, we held additional meetings with the working groups participating in the project, resulting in the identification of collaborative research questions and datasets. We published several datasets annotated with Linked Open Data and solicited feedback from data authors and users (most notably, those involved in the large-scale Anatolia Working Group). We also established collaborations with publishers, with the aim to situate open data dissemination more firmly within professional scholarly communication practices. Technical improvements during the second phase of the project included the "soft launch" of map tiles to enable faceted search on geospatial metadata in Open Context.

The project's second year involved additional meetings with working groups, publication of datasets, evaluation, continued improvements in data quality, and publication of our research activities. We also made significant progress in improving our software and Linked Open Data standards compliance and implementations. Major technical developments to Open Context during this phase centered on a complete rebuilding of Open Context using updated technologies that are easier to maintain, adapt, scale, and offer linked data. During this period we migrated Open Context's data to a new system built on the Djano (Python) framework, with a Postgres backend database and updated Apache Solr index. These changes made Open Context ready to scale on cloud computing environments over the next several months.

The project was extended into a third year due to additional funding from the Encyclopedia of Life (EOL), the J.M. Kaplan Fund, the German Archaeological Institute (DAI) and the Center for Hellenic Studies (CHS), which supported technical developments and the addition of a major new working group, the Anatolia Working Group (discussed below). During the third year of the project, we launched a new version of Open Context and finalized several publications related to this work.

Additional project activities involved collaboration with the DIPIR project (http://dipir.org), an IMLS-funded initiative to study data-reuse among researchers in different disciplines (see Faniel et al. 2013). The DIPIR team conducted qualitative research on data-reuse among two of our project's working groups. This work resulted in several presentations and publications, as well as continuing collaborative work among our teams.

RESEARCH OUTCOMES: DATA PUBLISHING WORKFLOWS

Open Context's model of "data sharing as publishing" emphasizes a more formalized publication process, where editors and reviewers collaborate with data contributors to identify and resolve issues, annotate and align data to expected standards, and work together to improve the overall quality and intelligibility of data, including peer review (see Figure 1). While this approach is more labor-intensive, it does a better job of ensuring that the data can be understood and reused by others. This project aimed to demonstrate that formalized data dissemination (rather than simply putting data into a digital repository) is critical to enabling the exchange of information and replication of research results. We achieved this by observing first-hand data documentation, integration, and reuse among several groups of researchers. Below, we describe each group and their major outcomes.

Our study focused on organizing "working groups" of scholars with common research interests involving the analysis of shared data. These working groups included the Archaeometallurgy Working Group (sharing laboratory generated chemical analyses), the Archaeological Ceramics Working Group (sharing petrographic and neutron activation analysis results), the Archaeological Survey Working Group (sharing results of archaeological surface surveys), the Central Mediterranean Working Group (sharing results of archaeological excavations in Italy and Tunisia), and finally the Anatolia Working Group (sharing results of zooarchaeological data in Turkey). These working groups had mixed outcomes. While all made progress in achieving data dissemination and collaborative analysis results, only a few achieved all their goals during the period of this project. Below we summarize these results.

- **Successes**: Very few researchers we approached about participating in data publishing and collaborative analysis outright refused. Most had great enthusiasm for the project and its goals, and despite delays and time distractions, most researchers committed significant effort and demonstrated clear progress. The biggest success came from the Anatolia Working Group, where more than 30 researchers shared hard-earned zooarchaeological data documenting over a dozen sites in Turkey. This working group published an often-cited synthetic article in *PLOS ONE* that has been viewed more than 4,000 times since June 2014 (Arbuckle et al. 2014). We also described the work-flows and editorial processes behind the success of this working group in an article that won the "Best Paper Prize" at the 2014 Data Curation Conference (Kansa et al. 2014).
- **Challenges**: The biggest challenge centered time commitments faced by researchers. In all cases, researchers complained they had very little time to actually conduct research. Teaching obligations, graduate student advising, but especially onerous administrative burdens all got in the way of dedicating time and effort into preparing datasets for publication. In addition, the working groups that attempted to recruit outside participants during the project faced delays; the fact that open data sharing on this scale is still very novel in archaeology meant that it took time to develop shared expectations and feelings of trust. Because of these problems, most of the working groups only made datasets available in a slow and piecemeal manner, delaying collaborative analysis. These working groups are still making progress, but their final outcomes will not be delivered until after the period of this grant.

DISCUSSION: HOW THE ANATOLIA WORKING GROUP SUCCEEDED

As discussed above, the Anatolia Working Group achieved the greatest successes in data and collaborative analysis outcomes. The one dozen participants in the Anatolia working group published zooarchaeological datasets, analyzed different subsets of the integrated data, and presented their

results to the group. Participants provided feedback on the data editing and integration process, and the group had lengthy discussions about using data produced by others. This feedback was integral to improving the data publishing process to facilitate reuse and is helping to streamline the process for the other working groups in our NEH-funded study.

We cannot understate how useful this working group was for developing a better understanding of the challenges in publishing meaningful data. We credit the comparative success of the Anatolia Working Group to a number of factors. Because zooarchaeologists describe specimens using relatively standard taxonomic and anatomical terms, their datasets were easier to understand and align. Thus, we were able to move beyond the initial steps of data integration and focus on some of the more challenging aspects of working with multiple datasets. Below we outline some specific observations that have advanced our understanding of data publishing:

- **Workflow Challenges:** In preparing multiple large zooarchaeological datasets for publication and collaborative analysis, we gained valuable insights into costs, efforts, and workflow challenges. The most significant challenge we note thus far relates to decoding data. Data are often collected in coded form for expedience in data entry. Unfortunately, codes are not always well documented or consistently applied. This creates tremendous problems in decoding (see more discussion in this blog post: http://ux.opencontext.org/blog/2013/02/26/decoding-data-a-view-from-the-trenches/).

- **Inappropriate Tools at Data Creation:** We have learned that even if a discipline shares a common methodological standard, inappropriate information management tools can stymie use of data created according to that standard. For example, the zooarchaeologists involved in the Anatolia Working Group all shared a common standard for documenting tooth-wear for making age-at-death estimates for sheep and goats. However, these data require somewhat complicated schema to adequately model. Such schema are difficult to express in a single table spreadsheet (such as Excel), typically used in data management. In practice, such tooth-wear data, though derived from common research methods, end up expressed as more-or-less free text in spreadsheets. Thus, the data, though "standardized", still required significant editorial work to process for publication and integration.

- **Cross Disciplinary Collaboration:** The Anatolia Working Group involved annotation with a sophisticated bio-informatics ontology, UBERON (http://uberon.org). This has helped us build bridges and collaborative ties with the bioinformatics community, and we are exploring future research involving UBERON-facilitated integration of genetic and epigenetic data with archaeological data. This helps illustrate how linked data approaches can facilitate cross-pollination of knowledge between humanities-oriented projects like Open Context and efforts in the natural sciences.

The Anatolia Working Group concluded with the publication of 14 datasets, as well as two publications describing the methodological and research results in detail (see *Audience, Dissemination, and Evaluation*). Since the publication of the datasets associated with this project, we know of two additional studies that have reused those datasets (in press publications by C. Cakilar[1,2]). The immediate successful

[1] Cakirlar C. (in press). Adaptation, Identity, and Innovation in Neolithic and Chalcolithic Western Anatolia (6800-3000 cal. BC): the Evidence from Mollusk Shells. *Quaternary International*.

reuse of data is helpful feedback for Open Context's data publishing process, and suggests that data clean-up, editing, description, and annotation give the datasets more meaning and make them more suitable for new analyses. The Anatolia Work Group has also been used in teaching. Ben Arbuckle (the working group chair) led a seminar at Cornell University on March 17, 2015, where he discussed the complexities of data integration and reuse, including issues of trust and data documentation (podcast available from the Cornell Institute of Archaeology and Material Studies: http://ciams.cornell.edu/2015/03/17/radiociams-ben-arbuckle-on-big-data-archaeology-and-neolithic-animal-economics/)

Learning from Data Publishing Successes and Challenges

The success of the Anatolia Working Group highlights how community, technical, and semantic context, all play key roles in the success of data sharing and collaborative analysis. Below we outline the roles played by these different factors:

- **Community:** Under Ben Arbuckle's leadership, participants in the Anatolia Working Group had already discussed and committed to data sharing and collaborative analysis years before our project was funded by NEH (and supplemented with EOL support). That is, the community had already established certain expectations for data sharing. These same expectations did not already exist in the other working groups, requiring more time and effort building such expectations. To help build community-wide expectations for sharing data, leaders of the other working groups shared their own datasets more or less in isolation as a proof of concept and principle. While none of the other working groups achieved the same level of community-wide buy-in as the Anatolia Working Group, they still created open datasets and in doing so, they helped promote wider data sharing in their own sub-disciplinary communities. The still forth-coming data publication projects resulting from these other working groups highlights the slower but still clear progress toward changing disciplinary cultures.
- **Technical:** The Anatolia Working Group also had a number of technical advantages that reduced the labor and time commitment costs to data sharing. The zooarchaeologists collected relatively clearly structured data (usually on single tables in Excel, sometimes more complex relational databases). Typically, they did not provide much supplemental media (images). On the other hand, the other working groups often created very media rich data. They faced greater challenges in personal data management, since not only do they need to create and curate databases, but they also need to manage directories of associated image files. This added complexity required much more time and effort to prepare data for submission to Open Context for publication. Unfortunately, there is no simple solution to this problem. Once entropy sets in and data become disorganized, it can be very time consuming and tedious to put datasets and media directories back into order. Better data management training and more deeply rooted professional expectations for data dissemination are needed to encourage better personal data management practices.

---

[2] Çakirlar, C., & Atici, L. (in press). Patterns of animal exploitation in western Turkey: from Palaeolithic molluscs to Byzantine elephants. In U. Albarella, H. Russ, K. Vickers, & S. Viner-Daniels (Eds.), *Oxford Handbook of Zooarchaeology*. Oxford: Oxford University Press.

- **Semantic:** Finally, the Anatolia Working Group benefited from a much more mature information ecosystem in terms of available ontologies and controlled vocabularies. Zooarchaeology has more consistent recording systems and research methods than other areas of archaeology. It can also draw upon mature ontologies and controlled vocabularies such as those offered by the EOL and Uberon. In contrast, material culture varies much more in place and time than biological species. Therefore, the archaeological community has much steeper challenges in building the "semantic infrastructure" needed for interoperability in material culture. Researchers focusing on material culture also have a harder time building a critical mass of related data. A given material culture dataset has narrower chronological and geographic relevance to other datasets. These factors mean that synthetic research outcomes based on the analysis and interpretation of aggregated material culture datasets require much longer time horizons. While our project demonstrated (with zooarchaeology) successful research outcomes with data integration, similar successes with material culture data will need more time. Achieving these longer-term goals requires pathways that give incremental progress. We demonstrated such incremental progress by showing how publishing even single datasets (in relative isolation) can still benefit research because even a single dataset can promote greater reproducibility, transparency, and trust in interpretations. Publications in *Internet Archaeology* (Thompson and Skaggs 2012), *Antiquity* (Grave et al. 2014), and the *Journal of Archaeological Science* (Grave et al. 2013) that reference data published by Open Context demonstrate the feasibility of an incremental path toward significant research outcomes with material culture data.

Additional Research Outcomes (The DIPIR Collaboration)

This project has benefited greatly from ongoing collaboration with the DIPIR project (http://dipir.org), an IMLS-funded initiative to study data-reuse among researchers in three different disciplines. With the agreement and consent of the participating researchers, we invited the DIPIR project to conduct qualitative research on data-reuse in this study. Previously, DIPIR only had access to retrospective accounts of how researchers reused data. We gave the DIPIR team unprecedented access to study, in "real time", the back-and-forth communications and problem solving between Open Context data editors and researchers as they collaborated on data decoding, cleaning, documentation, annotation, and analysis.

Technology Outcomes

During the grant period, we undertook a complete software rebuild of Open Context. The changes, including shifting programming languages from PHP to Python and integrating Postgres as a backend database, were aimed at better supporting the publication of Linked Open Data. Past successes in using data publishing and Linked Open Data annotation to encourage data reuse and syntheses in zooarchaeology (see Kansa et al. 2014; Arbuckle et al. 2014) motivated these technical investments. The revised version of Open Context makes "machine-readable" (easy to parse by software) data more easily available with updated and redesigned APIs. This facilitates interoperability by enabling users and their software agents to access data in formats that can be easily loaded into other databases and combined with other data, or can be easily manipulated by software for analysis and visualization. Specific software accomplishments include the following:

- *PeriodO Implementation:* The upgrade of Open Context enables more theoretically justifiable ways of organizing archaeological data. One of the most important aspects of this upgrade centers on modeling archaeological time periods. Chronological periods are key organizing principles in archaeology (Rabinowitz 2014). In order to offer theoretically appropriate ways of periodizing archaeological data, we have implemented PeriodO, another NEH-ODH funded project developing a "gazetteer" of time periods using Linked Open Data.

- *Mobile-Friendly Interface:* Technical developments included integration of the latest version of Bootstrap, an open source front-end development framework. We are continually deploying and testing interface features for the new version of open context, and these check as mobile-friendly by Google Developers (https://www.google.com/webmasters/tools/mobile-friendly/?url=http://opencontext.dainst.org/).

- *GeoJSON API:* Developments included updating the primary structured data representation for Open Context from an XML vocabulary to a Linked Open Data version of GeoJSON. Updates to Open Context's search API make all of Open Context's APIs consistent and much easier for third parties to use. Research applications of the new API include Shawn Graham's topic modeling of field notes published in Open Context used the new API: http://rpubs.com/shawngraham/79365. In addition, rOpenSci sponsored an R statistical package that uses our new API. This was built by Ben Marwick and Lincoln Mullen at a hackathon at GitHub headquarters in San Francisco (found here: https://github.com/ropensci/opencontext. The new Open Context API is documented here: http://opencontext.dainst.org/about/services.

The new developments in Open Context have enabled much greater levels of interoperability with a host of allied systems (see Table 1). This interoperability ensures that the data published by Open Context has a wide reach to related content on the Web, thus increasing its usefulness for research, teaching, and scholarly inquiry.

Improved Search and User Interface Design

Revisions to Open Context's faceted search focused on making much smaller and easier to maintain code and taking better advantage of mature open source software (Apache Solr). In addition to a tuned-up search interface, Open Context now has a GeoJSON-LD search API. However, for getting search results from Open Context, we are continuing to use the old but very useful Atom feed API, which has great utility in sharing lists of search results. Developments to Open Context use GeoJSON-LD as a common representation format for all Open Context data. In addition to making publicly-available, machine-readable data, Open Context reads GeoJSON-LD data to show records in its own interface. The open source Bootstrap libraries provide the layout, typography, styling, and various interactive features (drop-down lists, tabs, accordion boxes). The new grid layout in Open Context is mobile friendly and resizes well.

Emphasizing Linked Open Data

Since Open Context's inception nine years ago, we now have a better understanding of some of the key issues and requirements for managing this kind of scale and diversity of archaeological data. At the

same time, the archaeological "information ecosystem" has also grown. Our community has made great strides in sharing more and more interoperable data. This growth, both in the amount and sophistication of data and data users, has prompted us to make some significant revisions to Open Context. These revisions are aimed at getting more capability with less software because we have shifted to better models and abstractions.

International Collaboration

As documented by GitHub (https://github.com/ekansa/open-context-py/commits/master) the revision of Open Context has involved substantial effort, taking place in parallel with continued data publication. We prioritized software development to promote interoperability with the German Archaeological Institute (DAI), one of the leading archaeological research institutions in the globe. The DAI is providing mirror hosting and back-up for Open Context, and Open Context is developing data publication workflows for DAI content.

Lessons, Recommendations, and Future Directions

Continued Expansion of Open Data

In working with data authors over several years, we have found that the reservations researchers have about data publishing mainly stem from a lack of understanding of how data publishing works and the draw-backs on one-off, "siloed" systems. With regard to the latter point, we wanted to be careful not to discourage the creation of other data sharing infrastructure, since researcher informatics needs are diverse and developing new data sharing systems can advance the state of the art. Instead, we wanted to encourage better practice in terms of data interoperability, openness, persistent identity, version control and digital preservation. Encouraging better practice in other data sharing systems would still advance our NEH project goals, without casting data sharing as a zero-sum game, advancing one system at the expense of another.

Keeping the "Human" in the Humanities

Face-to-face meetings were invaluable to this project. Phone calls and emails were useful for formally communicating the project aims to potential participants and for dealing with the details of publishing individual datasets. However, because this project is highly collaborative, the in-person working group meetings were where the "big picture" could be fleshed out through active discussion by all participants. These meetings also established more trust among the working group participants—essentially putting a face to a dataset. This more personal relationship between researchers establishes trust and is one of the key elements underlying confidence in data reuse (Faniel et al. 2013). Finally, as demonstrated by the Anatolia Working Group, success in collaborative data dissemination and reuse requires long-term development of trust, shared expectations, and common research interests (in collaborative analysis). Ben Arbuckle's role in building this trust and consensus helped the Anatolia Working Group achieve great success.

Method, Theory, and Data Standards

Archaeology faces huge challenges in developing the needed semantic infrastructure that would enable data integration (at least for material culture studies). We should also note that archaeology has other methodological and theoretical concerns than efficient interoperability. Archaeology can be described as an artisanal craft (Shanks and McGuire 1996), and as such, many archaeologists reject attempts to "mass-produce" standardized and highly fungible data. The key need for the discipline is *not* to standardize what archaeologists say or cannot say about the past. Rather, we should aim for data management practices that make modeling and classification, including definition of new classification schemes, more formal and explicit. If archaeologists want to meaningfully reuse and compare datasets from multiple field projects, and if they do not want to accept standardized recording practices, then they must accept greater responsibility in formally and precisely documenting and modeling their own "customized" approaches to organizing data.

An increasing number of digital humanities projects are following this idea of "formalization over standardization", which forms the foundation of Linked Open Data. Formal annotation, rather than forced standardization, provides the flexibility needed to integrate content from many projects. For example, the new PeriodO project (http://perio.do) illustrates the value of publishing research-defined classification systems using computational formalism. (Shaw et al. 2015). PeriodO models the geographic and temporal scope of a period, including information about the authority that defined the period. It does not demand agreement where agreement does not exist, but still enables interoperability.

Looking forward, we realize with Open Context that we need to expand our role as a "data publisher". In addition to publishing datasets, Open Context needs to do more to publish the controlled vocabularies and models needed to relate different datasets together. This expansion of service will help meet a critical need in archaeology to create, in a "bottom-up" and more theoretically-grounded fashion, the semantic infrastructure needed to make sense of material culture data at large scales. The recent award of an NEH Research and Development Grant to the Alexandria Archive Institute (PR-234235, starting January 1, 2016) will provide the resources to develop these new publishing services.

Audience, Dissemination, and Evaluation

Audience

It is very difficult to count the overall number of users of Open Context. We have adopted strict privacy protections, based on the American Library Association's recommendations, so we do not use tracking tools like Google Analytics. Our service logs record on the order of 300 to 500 unique visitors per day, but it is difficult to know how many of these are software agents and how many are users. It is easier to understand publication and citation of Open Context rather than Web interactions. The publications and presentations listed below illustrate the diversity of audiences this work serves. These audiences include digital humanists, librarians, museum professionals, archaeological researchers, scholars conducting field work, professional archaeologists working in cultural resource management, heritage managers working in public and government contexts, and scholars and students working in archaeological specializations.

DISSEMINATION

A key goal of this project is to encourage greater professional recognition for data publication as an accepted and expected form of scholarly communication. To this end, project participants have broadcast this work as widely as possible through publications, conference presentations, forum discussions, and social media outlets.

Over three years, this project produced the following relevant products, including 14 publications, 37 conference presentations, and 19 data publications.

PUBLICATIONS

(* Indicates articles containing links to related datasets published in Open Context.)

*Arbuckle BS, Kansa SW, Kansa E, Orton D, Çakırlar C, Gourichon L, Atici L, Galik A, Marciniak A, Mulville J, Buitenhuis H, Carruthers D, De Cupere B, Demirergi A, Frame S, Helmer D, Martin L, Peters J, Pöllath N, Pawłowska K, Russell N, Twiss K, and Würtenberger D. 2014. Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals across Neolithic Turkey. *PLOS ONE* e99845. DOI: 10.1371/journal.pone.0099845.

Faniel, Ixchel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, and Elizabeth Yakel. 2013. The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse *JCDL 2013 Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*: 295-304. New York, NY. DOI: 10.1145/2467696.2467712

*Grave, Peter, and McNiven, Ian J. 2013. Geochemical provenience of 16th–19th century C.E. Asian ceramics from Torres Strait, northeast Australia. *Journal of Archaeological Science* 40:4538–4551.

*Grave, Peter, Lisa Kealhofer, Ben Marsh, Ulf-Dietrich Schoop, Jürgen Seeher, John W. Bennett, and Attila Stopic. 2014. Ceramics, trade, provenience and geology: Cyprus in the Late Bronze Age. *Antiquity* 88: 1180-1200. doi:10.1017/S0003598X0011539X.

Kansa, Eric. 2012. Openness and archaeology's information ecosystem. *World Archaeology* 44:4 pp. 498-522. (Open access pre-print: http://alexandriaarchive.org/blog/wp-content/uploads/2012/Kansa-Open-Archaeology-Self-Archive-Draft.pdf)

Kansa, Eric. 2014. The Need to Humanize Open Science. In *Issues in Open Research Data*, edited by Samuel A. Moore, 31-58. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/ban

Kansa, Eric. 2015. Contextualizing Digital Data as Scholarship in Eastern Mediterranean Archaeology. *CHS Research Bulletin* 3(2). [Available open access: http://www.chs-fellows.org/2015/08/03/contextualizing-digital-data-as-scholarship-in-eastern-mediterranean-archaeology/)

Kansa, Eric C. 2015. Reimagining Archaeological Publication for the 21st Century. In A. Traviglia, ed. *Across Space and Time: Papers from the 41st Conference on Computer Applications and*

Quantitative Methods in Archaeology, Perth, 25-28 March 2013. pp. 367-378. Amsterdam University Press.

Kansa, Eric C. and Sarah Whitcher Kansa. 2013. We All Know That a 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1:88-98.
[Note: Five responses to the paper and a rejoinder by the authors focus in particular on the sustainability of open data.] The paper is available at:
http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/journal_of_eastern_mediterranean_archaeology_and_heritage_studies/v001/1.1.kansa01.html

Kansa, Eric, Sarah Whitcher Kansa, and Benjamin Arbuckle. 2014. Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology. *International Journal of Digital Curation* 9(1):57-70. DOI: 10.2218/ijdc.v9i1.301 [Note: Winner of the "Best Paper Award" at the 2014 IDCC conference (San Francisco)]

Kansa, Eric C., Sarah Whitcher Kansa, and Lynne Goldstein. 2013. On Ethics, Sustainability, and Open Access in Archaeology. *The SAA Archaeological Record* 13 (4):15-22. [Available open access]

Kansa, Sarah Whitcher and Carrie Dennett. 2015. Exploring Open Access for SAA Publications. *The SAA Archaeological Record* 15 (2): 5-8. [Available Open Access]

Kansa, Sarah Whitcher and Eric Kansa. 2015 "Reflections on a Road Less Traveled: Alt-Ac Archaeology." *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 3(3): 293-298. DOI:10.5325/jeasmedarcherstu.3.3.0293 [Article] [Open Access Preprint]

Kansa, Sarah Whitcher and Eric Kansa. 2014. Data Publishing and Archaeology's Information Ecosystem. *Near Eastern Archaeology* 77(3): 223-227.

*Thompson, Christine and Sheldon Skaggs. 2012. King Solmon's Silver?: Southern Phoenician Hacksilber Hoards and the Location of Tarshish. *Internet Archaeology*.

CONFERENCE PRESENTATIONS & WORKSHOPS

2012

1. American Schools of Oriental Research (Chicago, IL). Paper: *From Data to Knowledge: Organization, Publication, and Research Outcomes* (by Sarah W. Kansa), in the session "Topics in Cyberinfrastructure, Digital Humanities, and Near Eastern Archaeology."
2. PNC Conference, New Horizons: Information Technology Connecting Culture, Community, Time, and Place (Berkeley, CA). Paper: Applying Linked Open Data: Refining a Model of "Data Sharing as Publication" (by Eric Kansa) (http://www.pnclink.org/pnc2012/english/index.html)
3. GIS Day, Purdue University (Purdue, IN). Keynote presentation: *Using the Web to Situate Archaeology in Place, Time, and Community* (by Eric Kansa) (https://stemedhub.org/groups/2012gisday/gis_day_college_program)

2013

1. Computer Applications in Archaeology annual conference (Perth, Australia). Keynote presentation: *Reimagining Archaeological Publication for the 21st Century* (by Eric Kansa) (http://www.caa2013.org/drupal/speakers)
2. Opening the Past conference (Pisa, Italy). Keynote presentation: *Open Access in Archaeology* (by Eric Kansa) (http://www.mappaproject.org/wp-content/uploads/2011/08/Pisa-OpenAccess-Kansa-FINAL.pdf)
3. Joint Conference on Digital Libraries (Indianapolis, IN). Presentation: *The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse* (co-authored by Ixchel Faniel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, Elizabeth Yakel)
4. Harvard Digital Futures Consortium (Cambridge, MA). Inaugural presentation: *A More Open Future for the Past* (by Eric Kansa) (http://df.darthcrimson.org/). [Note: A September 17 article in the Harvard Gazette presents highlights of the talk: http://news.harvard.edu/gazette/story/2013/09/the-modern-opens-the-past/; the talk is also online: http://www.youtube.com/watch?v=trJPbqEKEhE&feature=player_embedded]
5. American Schools of Oriental Research (Baltimore, MD). Working group meeting.
6. American Schools of Oriental Research (Baltimore, MD). Presentation: *'Big Data' and Collaborative Research in Zooarchaeology* (co-authored by Sarah Kansa and Benjamin Arbuckle)
7. American Schools of Oriental Research (Baltimore, MD). Presentation: *Data Management as Scholarly Communications in Near Eastern Archaeology* (by Eric Kansa)
8. American Schools of Oriental Research (Baltimore, MD). Presentation: *From the Ground Up: The Construction of the Cisjordan Corpus and Its Data Set As A Platform for Frequency Specific Metallic Sequencing and Data Sharing As Publication* (by Christine Thompson)
9. American Schools of Oriental Research (Baltimore, MD). Presentation: *Late Bronze Age Pot Marks and Provenience: The Problem of 'Hidden Data'* (by Michael Sugerman)
10. American Schools of Oriental Research (Baltimore, MD). Forum: *Scholarship in the 21st Century* (Eric Kansa, panelist)
11. UC Berkeley Data Science Institute dedicatory conference (Berkeley, CA). Poster Presentation / Live Demo: *Open Context and Data Sharing as Publication* (by Eric Kansa)

2014

1. Archaeological Institute of America conference (Chicago, IL). Forum: *The Future of Archaeological Publishing* (Eric Kansa, panelist)
2. UC Berkeley D-Lab Data Science Institute (Berkeley, CA). Presentation / Live Demo: *Models for Publishing Research Data* (by Eric Kansa)
3. International Digital Curation Conference (San Francisco, CA). Presentation: *Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology* (by Eric Kansa, Sarah W. Kansa, and Benjamin Arbuckle).
4. Innovating Communication in Scholarship (ICIS) conference on the theme "Publish or Perish – the Future of Academic Publishing and Careers" (Davis, CA). Forum: *Beyond Journals & New Forms of Digital Publishing* (Eric Kansa, panelist). [See the blog post with links to conference videos: http://ux.opencontext.org/blog/2014/03/05/publish-or-perish-2014-resurrected-online/]

5. University of Pennsylvania, Digital Humanities Forum Symposium (Philadelphia, PA). Forum: *Research on a Global Scale: The Radical Potential of Linked Open Data* (Eric Kansa, panelist)
6. Forum: *Archaeological Publishing in the 21st Century*. Sarah Kansa was a panelist in the opening night presidential forum, focusing on new forms of publishing, including data and open access. (Society for American Archaeology conference, Austin, April 2014)
7. Society for American Archaeology conference (Austin, TX). Presentation: *Documenting and Disseminating Zooarchaeological Data in the Digital Age* (Sarah Whitcher Kansa, Levent Atici, Richard H. Meadow, and Eric C. Kansa)
8. University Libraries' Data Speaker Series, Washington University (St. Louis, MO). Invited Presentation: *Linked Data: Publishing to the Web of Data in Archaeology* (by Eric Kansa)
9. International Council for Archaeozoology conference (San Rafael, Argentina). Presentation: *Data as Professional Practice in Zooarchaeology* (by Sarah W. Kansa)
10. International Council for Archaeozoology conference (San Rafael, Argentina).Workshop: *Digital Data: Collection, Organization, and Dissemination* (led by Sarah W. Kansa)
11. 8Th Congress of Archaeology in Berlin (Berlin, Germany). Presentation: *Linked Data, Publication, and the Life Cycle of Archaeological Information* (by Eric Kansa)
12. American Schools of Oriental Research (San Diego, CA). Presentation: *Published but Perished Anyway? Moving Archaeology toward Open, Collaborative and Data Intensive Research* (by Eric Kansa)
13. American Schools of Oriental Research (San Diego, CA). Presentation: *Near Eastern Zooarchaeology – Is there a Future for our Data?* (by Sarah W. Kansa)

2015

1. Society for American Archaeology conference (San Francisco, CA). Presentation: *Academic Freedom, Data, and Job Performance in the Panopticon* (by Eric Kansa)
2. Society for American Archaeology conference (San Francisco, CA). Forum: *Evaluating and Rewarding 21st Century Archaeological Scholarship* (Sarah W. Kansa, organizer).
3. Society for American Archaeology conference (San Francisco, CA). Forum: *Diverse Digital Archaeologies* (Sarah W. Kansa, panelist).
4. Center for Hellenic Studies (Washington DC). Presentation: *Contextualizing Digital Data as Scholarship in Eastern Mediterranean Archaeology* (by Eric Kansa)
5. Linking the Middle Ages Workshop (Austin, TX): Invited presentation: *Scholarly Communications and Linked Open Data in Archaeology* (by Eric Kansa)
6. Dryad Community Meeting (Washington DC, May 2015): Invited Panelist (Eric Kansa)
7. Institute on Digital Archaeology Method & Practice (Lansing, MI). Eric Kansa was a faculty member for the NEH-funded institute organized by Michigan State University's Department of Anthropology and MATRIX: The Center for Digital Humanities and Social Sciences. The Institute brought together archaeologists and closely associated scholars interested in developing critical, hands-on skills in digital method and practice.
8. 2015 Chacmool Conference (Calgary, AL). Presentation: *Data as Professional Practice in Archaeology* (by Sarah W. Kansa)
9. American Schools of Oriental Research, Atlanta, GA). Presentation: *Connecting the Dots: Data Publishing and Synthesis for the Archaeology of Cyprus* (by Eric Kansa, Sarah Whitcher Kansa, and William Caraher)

10. American Schools of Oriental Research, Atlanta, GA). Presentation: *Research Impacts of Linked and Open Archaeological Data: Case Studies from the Eastern Mediterranean* (by Eric Kansa)

Social Media

Blog post (Feb 2013): "Decoding Data- A View from the Trenches" (http://ux.opencontext.org/blog/2013/02/26/decoding-data-a-view-from-the-trenches/)

Blog post (May 2013): "Lessons in Data Reuse, Integration, and Publication" (http://ux.opencontext.org/blog/2013/05/02/lessons-in-data-reuse-integration-and-publication/)

Blog post (Dec 2013): *It's the Neoliberalism, Stupid: Why Open Access / Data / Science is not Enough* (http://www.alexandriaarchive.org/blog/?p=931), by Eric Kansa. This post was widely discussed and reposted in a number of venues, including the Impact Blog of the London School of Economics.

Blog post (June 2014): "Zooarchaeology of Neolithic Anatolia: Research Outcomes from Large-Scale Data Integration with Open Context" (http://ux.opencontext.org/2014/06/16/zooarchaeology-of-neolithic-anatolia-research-outcomes-from-large-scale-data-integration-with-open-context/)

Blog post (Sept 2014): "Research outcomes of multi-author collaboration using open data" (http://ux.opencontext.org/2014/09/30/research-outcomes-of-multi-author-collaboration-using-open-data/)

April 2015: Eric Kansa's chapter in Issues in Open Research Data (2014, S. Moore, editor) is highlighted in a review from the Virginia Tech Open Research Network (https://blogs.lt.vt.edu/openvt/2015/04/15/book-review-issues-in-open-research-data/)

DATA PUBLICATIONS

The following datasets have been published with the support (in whole or in part) of our NEH Digital Humanities Implementation grant. All of the data publications listed below are associated with conventional publications that draw on the data.

Project	Data Contributor / Key Project Participant	Project DOI
Barçin Höyük	Alfred Galik	http://dx.doi.org/10.6078/M78G8HM0
Çatalhöyük (East and West Mounds)	David Orton	http://dx.doi.org/10.6078/M7G15XSF
Çatalhöyük (TP area)	Arek Marciniak	http://dx.doi.org/10.6078/M7B8562H
Cukurici Hoyuk	Alfred Galik	http://dx.doi.org/10.6078/M7D798BQ
Domuztepe	Sarah Whitcher Kansa	http://dx.doi.org/10.6078/M7SB43PP
Erbaba Höyük	Ben Arbuckle	http://dx.doi.org/10.6078/M70Z715B
Ilipinar	Hijlke Buitenhuis	http://dx.doi.org/10.6078/M76H4FB
Karain Cave	Levent Atici	http://dx.doi.org/10.6078/M7CC0XMT
Kösk Höyük	Ben Arbuckle	http://dx.doi.org/10.6078/M74Q7RW8
Okuzini Cave	Levent Atici	http://dx.doi.org/10.6078/M73X84KX
Pinarbasi (1994)	Denise Carruthers	http://dx.doi.org/10.6078/M7X34VD1
Suberde	Ben Arbuckle	http://dx.doi.org/10.6078/M70Z715B
Ulucak Höyük	Canan Cakirlar	http://dx.doi.org/10.6078/M7KS6PHV
Hacksilber Project	Christine Thompson	http://dx.doi.org/10.6078/M74M92GB
Ceramics, Trade, Provenience and Geology: Cyprus in the Late Bronze Age	Peter Grave	http://dx.doi.org/10.6078/M70V89RM
Murlo (Poggio Civitate)	Anthony Tuck	http://opencontext.org/projects/DF043419-F23B-41DA-7E4D-EE52AF22F92F
Animal Remains from Forcello	Angela Trentacoste	http://opencontext.org/projects/bdc6cb19-e739-404b-8b4f-eeeaae3b8aae
Asian Stoneware Jars	Peter Grave	http://opencontext.org/projects/4B16F48E-6F5D-41E0-F568-FCE64BE6D3FA
Pyla-Koutsopetria Archaeological Project	David K. Pettegrew, William R. Caraher, R. Scott Moore	http://opencontext.org/projects/3F6DCD13-A476-488E-ED10-47D25513FCB2

EVALUATION AND IMPACT

USER FEEDBACK

Open Context's development follows an iterative process of design and user feedback. The project has grown organically in this way, based on the technical and professional needs expressed to us by data authors and data users. In response to researcher needs, Open Context has increasingly formalized aspects of its data dissemination services and now emphasizes a model of data sharing as (formal) publication. An important aspect of this change is the implementation of formal controlled vocabularies and ontologies to annotate publish datasets so that data are easier to discover and reuse.

We solicit feedback from data authors and data users on the data publishing process, interface, usability, and implementation of formal controlled vocabularies and ontologies to annotate publish datasets. Recording metrics for evaluation is difficult because Open Context's privacy policy ensures that we do not track user activities. Furthermore, metrics and "altMetrics" (alternative, mainly Web-

based metrics) seem to be a problematic way to evaluate significance and impact. Metrics lead to "apples versus oranges" counting problems. For example, the Çatalhöyük Zooarchaeology project (with 127,000 items) currently has 705 project views and 95,250 views of individual records. In contrast, though it is much smaller and lower-profile, the Murlo Excavations project (with 16,000 items) has 3,588 project views and 133,000 views of individual records. The reason for more activity around Murlo's individual records is due to the nature of the data, where most of the Murlo project data describes individual objects (with some 25,000 images to browse). Since the Çatalhöyük project centers on zooarchaeology, users are interested in downloads of data tables rather than views of individual bone records. Thus, while the Çatalhöyük Zooarchaeology data is best experienced through download, the Murlo dataset is best experienced through browsing. These different kinds of activities lead to different kinds of metrics. We worry that such nuance and contextual factors maybe lost in university bureaucracies seeking "objective" performance metrics. We raised some of these issues in a recent and widely-discussed blog post, reposted by the London School of Economics (http://blogs.lse.ac.uk/impactofsocialsciences/2014/01/27/its-the-neoliberalism-stupid-kansa).

Long-Term Impacts and Continuation of the Project

Honors & Awards

- Honor/Award: Eric Kansa discussed the potential of Linked Open Data in a panel discussion at the White House in June 2013, where he was honored as a Champion of Change in Open Science for his work with Open Context and open data (http://www.youtube.com/watch?v=a26cEwbyMGQ; at 0:25:55). For more information, visit the Open Context blog post: http://ux.opencontext.org/blog/2013/06/20/white-house-honors-contributions-to-open-science/
- We received the "Best Paper Award" from the 2014 IDCC conference for our presentation "Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology" (published in the *International Journal of Digital Curation* 9(1):57-70. DOI: 10.2218/ijdc.v9i1.301)
- In January 2016, the Archaeological Institute of America will present Open Context with their 2016 "Award for Outstanding Work in Digital Archaeology".

Other Impacts

- Starting in late 2015, Open Context is tracked as an academic publisher in Google Scholar
- Google Scholar reports 25 different publications that reference Open Context in just the past year (2015). Some of the articles discuss Open Context's data publishing model, while others reference datasets published by Open Context. These references highlight Open Context's impact in the advancing the practice of research data management as well as the impact of data published by Open Context. (see search results here: https://scholar.google.com/scholar?as_ylo=2015&q=%22http://opencontext.org%22)
- In March 2015, the German Archaeological Institute (DAI) deployed a mirror site for Open Context hosting and back-up (http://opencontext.dainst.org/). The DAI is the world's largest sponsor of archaeological research and has great prestige and recognition globally.

- In October 2014, the Society for American Archaeology publications program issued revised instructions for authors that includes guides on the use and citation of Open Context for data: http://www.saa.org/portals/0/saa/publications/styleguide/styleguide_final_813.pdf#page=35
- We are establishing relationships with an increasing number of publishers, as part of our efforts to situate open data dissemination more firmly within professional scholarly communication practices. These collaborations aim to help establish workflows, policies, and guidelines for authors publishing digital content in conjunction with their conventional publications. Collaborators include the *Journal of Open Archaeology Data*, *Internet Archaeology,* Lockwood Press, University Press of Colorado, and the Cotsen Institute of Archaeology Press (UCLA).

REFERENCES

Arbuckle BS, Kansa SW, Kansa E, Orton D, Çakırlar C, Gourichon L, Atici L, Galik A, Marciniak A, Mulville J, Buitenhuis H, Carruthers D, De Cupere B, Demirergi A, Frame S, Helmer D, Martin L, Peters J, Pöllath N, Pawłowska K, Russell N, Twiss K, and Würtenberger D. 2014. Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals across Neolithic Turkey. *PLOS ONE* e99845. DOI: 10.1371/journal.pone.0099845.

Faniel, Ixchel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, and Elizabeth Yakel. 2013. The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse *JCDL 2013 Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*: 295-304. New York, NY. DOI: 10.1145/2467696.2467712

Grave, Peter, and McNiven, Ian J. 2013. Geochemical provenience of 16th–19th century C.E. Asian ceramics from Torres Strait, northeast Australia. *Journal of Archaeological Science* 40:4538–4551.

Grave, Peter, Lisa Kealhofer, Ben Marsh, Ulf-Dietrich Schoop, Jürgen Seeher, John W. Bennett, and Attila Stopic. 2014. Ceramics, trade, provenience and geology: Cyprus in the Late Bronze Age. *Antiquity* 88: 1180-1200. doi:10.1017/S0003598X0011539X.

Kansa, Eric. 2012. Openness and archaeology's information ecosystem. *World Archaeology* 44:4 pp. 498-522. (Open access pre-print: http://alexandriaarchive.org/blog/wp-content/uploads/2012/Kansa-Open-Archaeology-Self-Archive-Draft.pdf)

Kansa, Eric. 2014. The Need to Humanize Open Science. In *Issues in Open Research Data*, edited by Samuel A. Moore, 31-58. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/ban

Kansa, Eric. 2015. Contextualizing Digital Data as Scholarship in Eastern Mediterranean Archaeology. *CHS Research Bulletin* 3(2). [Available open access: http://www.chs-fellows.org/2015/08/03/contextualizing-digital-data-as-scholarship-in-eastern-mediterranean-archaeology/)

Kansa, Eric C. 2015. Reimagining Archaeological Publication for the 21st Century. In A. Traviglia, ed. Across Space and Time: *Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology, Perth, 25-28 March 2013*. pp. 367-378. Amsterdam University Press.

Kansa, Eric C. and Sarah Whitcher Kansa. 2013. We All Know That a 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1:88-98.

Kansa, Eric, Sarah Whitcher Kansa, and Benjamin Arbuckle. 2014. Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology. *International Journal of Digital Curation* 9(1):57-70. DOI: 10.2218/ijdc.v9i1.301

Kansa, Eric C., Sarah Whitcher Kansa, and Lynne Goldstein. 2013. On Ethics, Sustainability, and Open Access in Archaeology. *The SAA Archaeological Record* 13 (4):15-22. [Available open access]

Kansa, Sarah Whitcher and Carrie Dennett. 2015. Exploring Open Access for SAA Publications. *The SAA Archaeological Record* 15 (2): 5-8. [Available Open Access]

Kansa, Sarah Whitcher and Eric Kansa. 2015 "Reflections on a Road Less Traveled: Alt-Ac Archaeology." *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 3(3): 293-298. DOI:10.5325/jeasmedarcherstu.3.3.0293 [Open Access Preprint]

Kansa, Sarah Whitcher and Eric Kansa. 2014. Data Publishing and Archaeology's Information Ecosystem. *Near Eastern Archaeology* 77(3): 223-227.

Rabinowitz, A., (2014). It's about time: historical periodization and Linked Ancient World Data. *ISAW Papers* 7(22). http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/rabinowitz/

Shanks, Michael and Randall H. McGuire. 1996. The Craft of Archaeology. *American Antiquity* 61(1):75-88.

Shaw, Ryan, Adam Rabinowitz, Patrick Golden, and Eric Kansa. 2015. A Sharing-Oriented Design Strategy for Networked Knowledge Organization Systems. *International Journal on Digital Libraries*. DOI:10.1007/s00799-015-0164-0 Preprint at https://www.researchgate.net/publication/280529967.

Thompson, Christine and Sheldon Skaggs. 2012. King Solmon's Silver?: Southern Phoenician Hacksilber Hoards and the Location of Tarshish. *Internet Archaeology*.

APPENDIX

Table 1: Interoperability Examples

System / Project	Example	Description
California Digital Library, Merritt Digital Repository	https://merritt.cdlib.org/m/ucb_open_context	Merritt uses Open Context APIs to accession content for long-term archival preservation.
GitHub	https://github.com/ekansa?tab=repositories	Open Context uses GitHub, a commercial service, for short-term version control and sharing of data.
tDAR (Digital Antiquity)	http://opencontext.dainst.org/sets/United+States?proj=52-digital-index-of-north-american-archaeology-dinaa&dc-subject=tdar*	Open Context connects to tDAR with APIs to cross-reference archaeological site records in the DINAA project with metadata in tDAR.
Intellectual Property Issues in Cultural Heritage (IPinCH)	http://opencontext.org/about/intellectual-property	Open Context connects with the IPinCH project, primarily not for technical interoperability, but rather to link with a community of indigenous and academic scholars to guide ethical approaches to data sharing and data privacy protections.
Pelagios	http://pelagios.dme.ait.ac.at/api/datasets/739128af1e9613de03706a57cef46976	Open Context implemented geospatial annotation APIs used by Pelagios, a large network of information systems primarily focused on the ancient Mediterranean.
PeriodO	http://perio.do/	Open Context provided time period data to the PeriodO project and is implementing the PeriodO data model to relate Open Context published data with data shared by a several other institutions.
Pleiades	http://pleiades.stoa.org/	Open Context uses Pleiades place identifies to link data with other systems on geospatial criteria.
Arachne	See example ('Arachne Comparanda'): http://opencontext.dainst.org/types/252A30E2-3F6C-4BB8-1148-FD2D27436185	Open Context links to Arachne, a major digital collection hosted by the German Archaeological Institute to relate material published by Open Context with relevant comparanda.
Çatalhöyük Living Archive	See example: http://opencontext.dainst.org/subjects/4892D0EB-684B-47E9-D810-F056B8F3DFF1	Open Context uses APIs from the relevant data in the Çatalhöyük Living Archive, an experimental project hosted by Stanford University exploring theoretical issues in interpreting excavation data.
PIDBA & EWHADP	http://opencontext.dainst.org/sets/United+States?linked=dinaa-cross-ref&proj=52-digital-index-of-north-american-archaeology-dinaa	The Paleo-Indian Database of the Americas (PIDBA) and the Eastern Woodlands Household Archaeology Data Project (EWHADP) are two specialized databases reference Open Context hosted URIs for DINAA project (American archeological sites).
Pleiades	http://pleiades.stoa.org/	Open Context uses Pleaides place identifies to link data with other systems on geospatial criteria.
Statistical Analysis and Visualization for Open Context	https://github.com/ropensci/opencontext	This software connects with the Open Context API to get data for statistical analysis and visualization using Rstat. Rstat is a popular programming language for numeric analysis in the sciences, social sciences and humanities. It is often used by advocates of "reproducible research" that seek to make both data and analysis software used in research studies open for inspection and reuse.

Figure 1: Open Context's Data Publishing Workflow, showing the movement of content from private to public spheres and the key players involved in each stage.

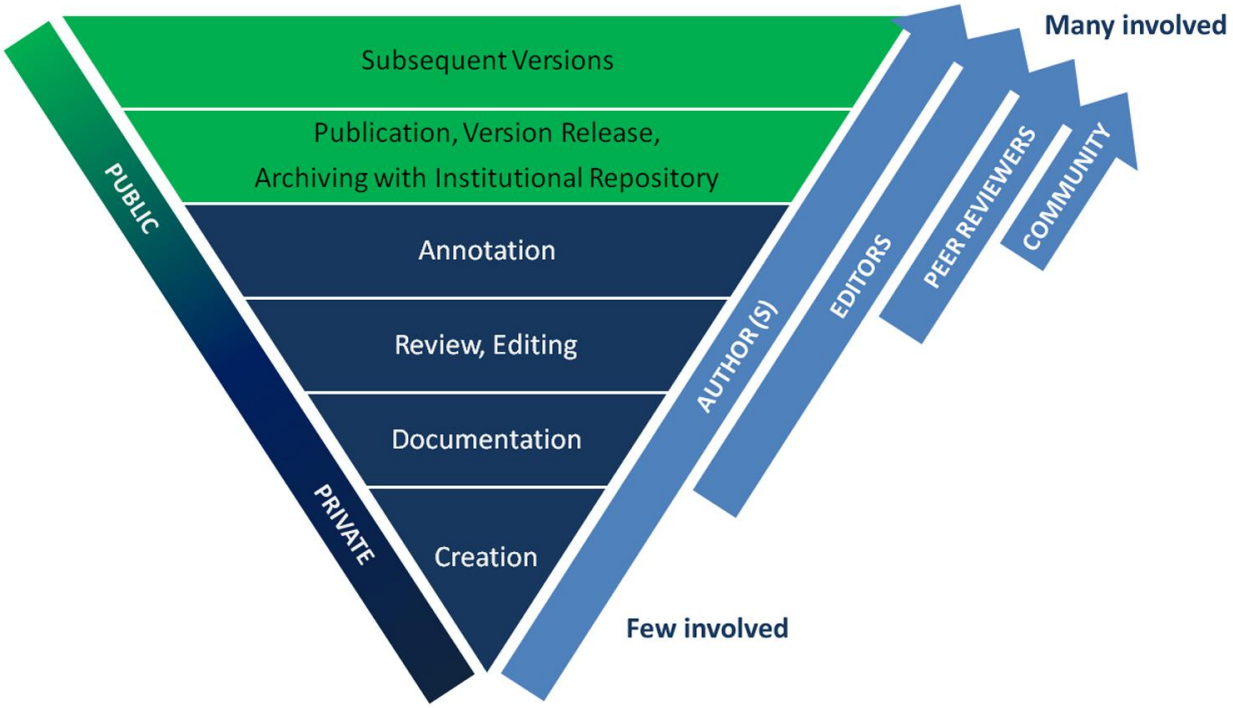


Figure 2: Open Context's new homepage, showing the global distribution of its data publications.

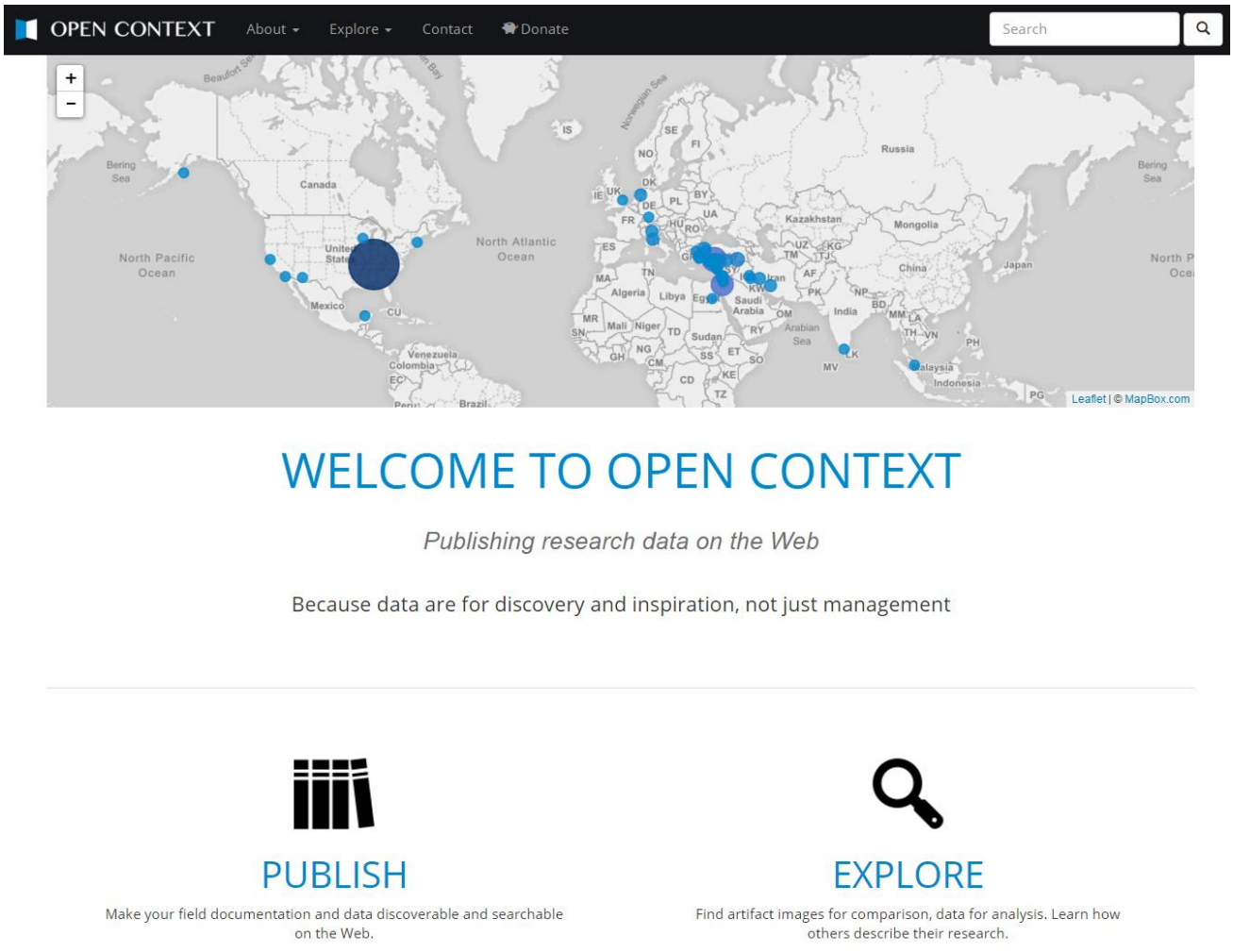


Figure 3: The new version of Open Context, showing the "project overview" for a data publication for one of this project's working groups. Including project images in the banner, a suggested citation, editorial status, and clear licensing give the data publication a more formal and professional "feel", while annotation and Linked Open Data provide flexibility for integration and reuse.

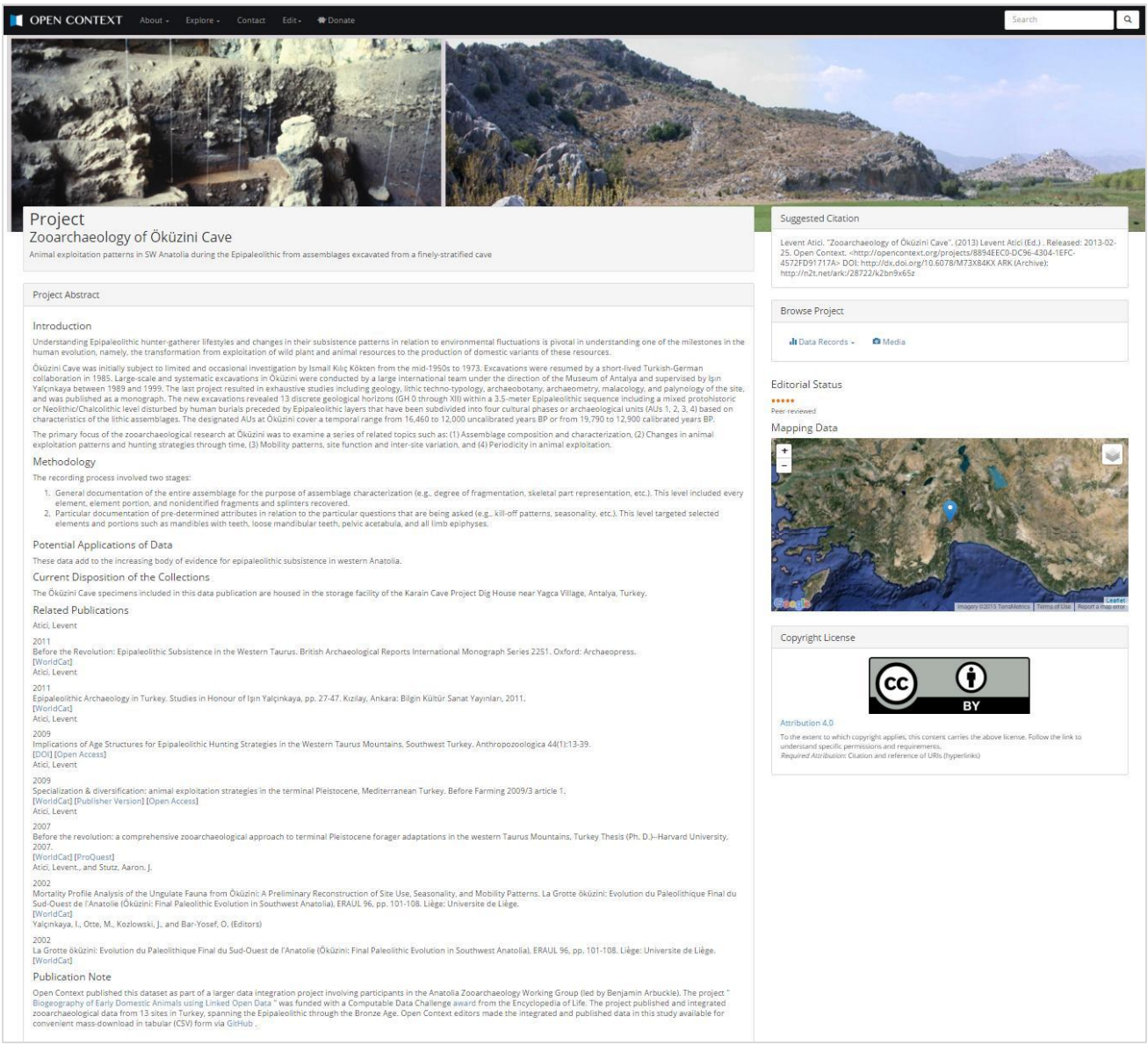
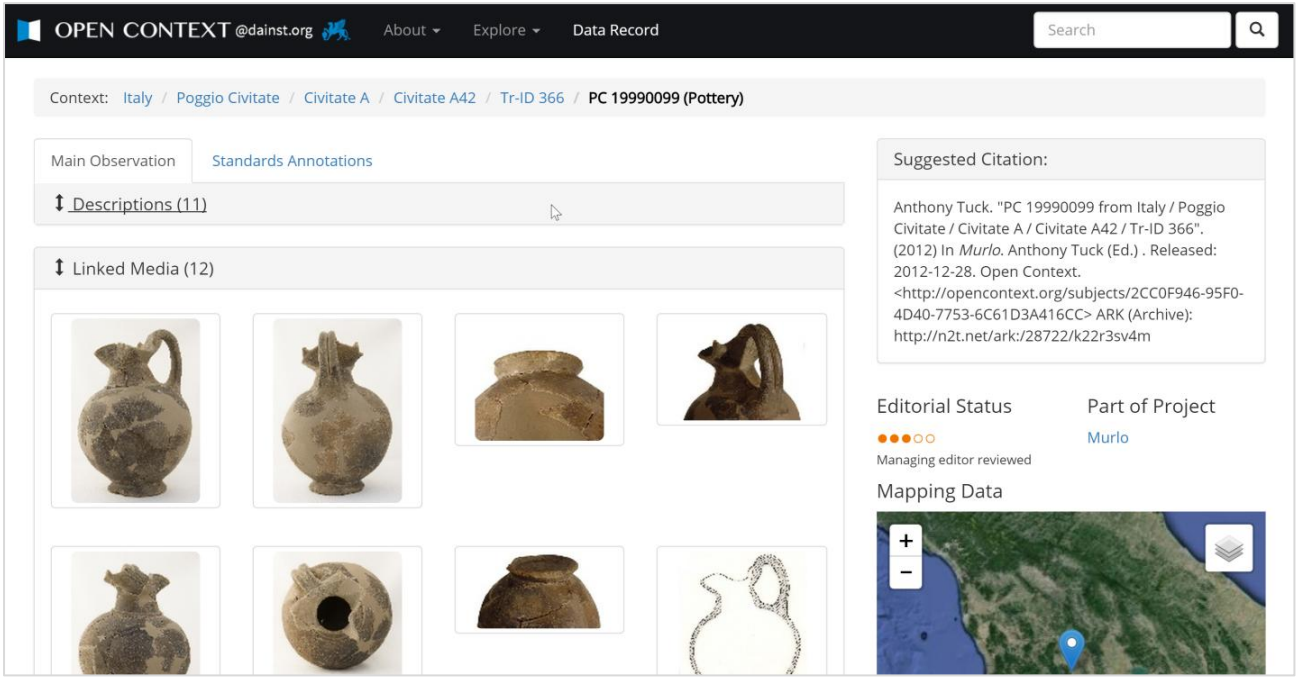


Figure 4: These two examples of Open Context's use of Linked Open Data show linking pottery types from Murlo, Italy with comparanda from another digital collection (Arachne). The display of comparanda from another system is enabled by the fact that Open Context and a growing number of other projects on the Web use Linked Open Data approaches (where related content from across the Web can be associated easily using unique web identifiers).



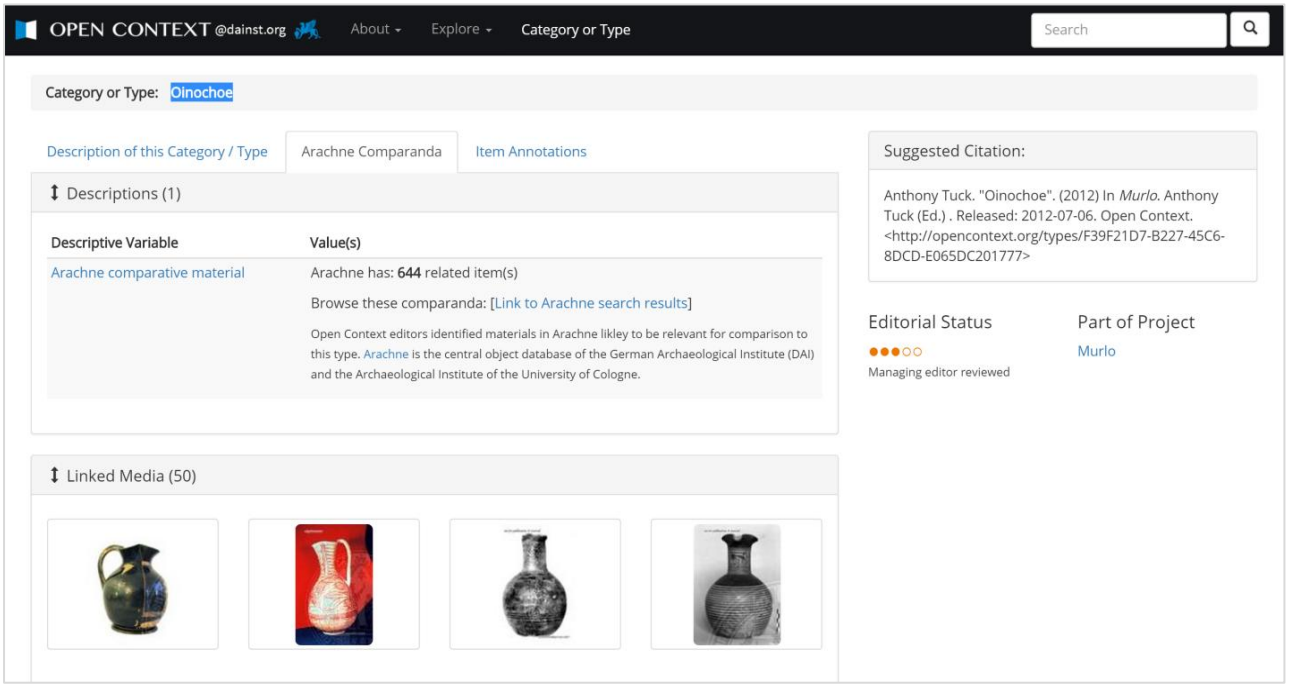


Figure 5a: View of a data record in the old version of Open Context. In the old view, the descriptive categories were all shown on the same page, regardless of the amount of content. In this example, the content continues down the page a long way, requiring the user to continue scrolling down. The record used here as an example is so large that the full record cannot be shown in this document.

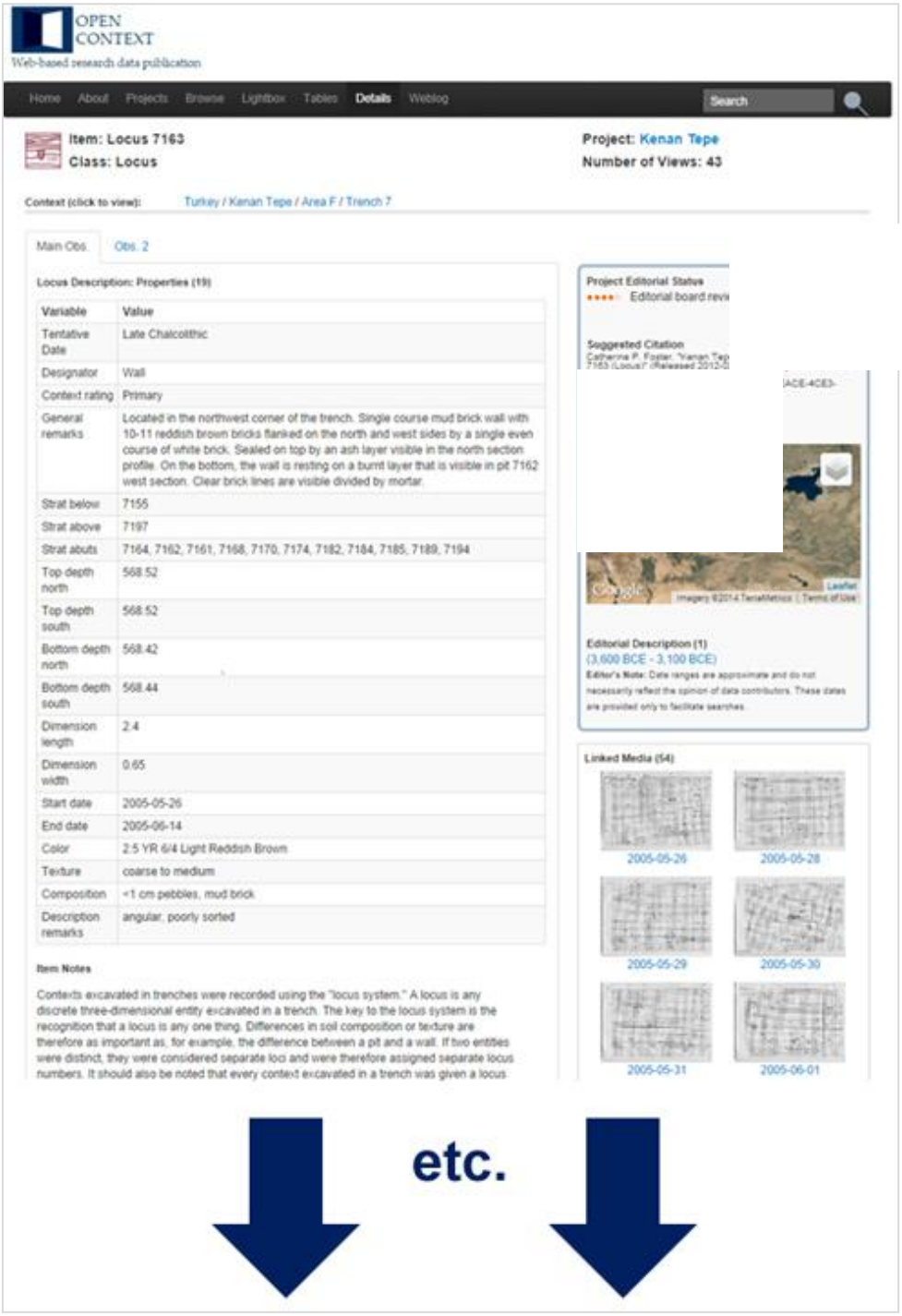


Figure 5b: View of the same data record in the new version of Open Context. In the new display, the descriptive categories are still all shown on the same page; however, the categories are collapsed to facilitate viewing. In this figure, all tabs are collapsed and each tab shows a count of how many items its section contains.

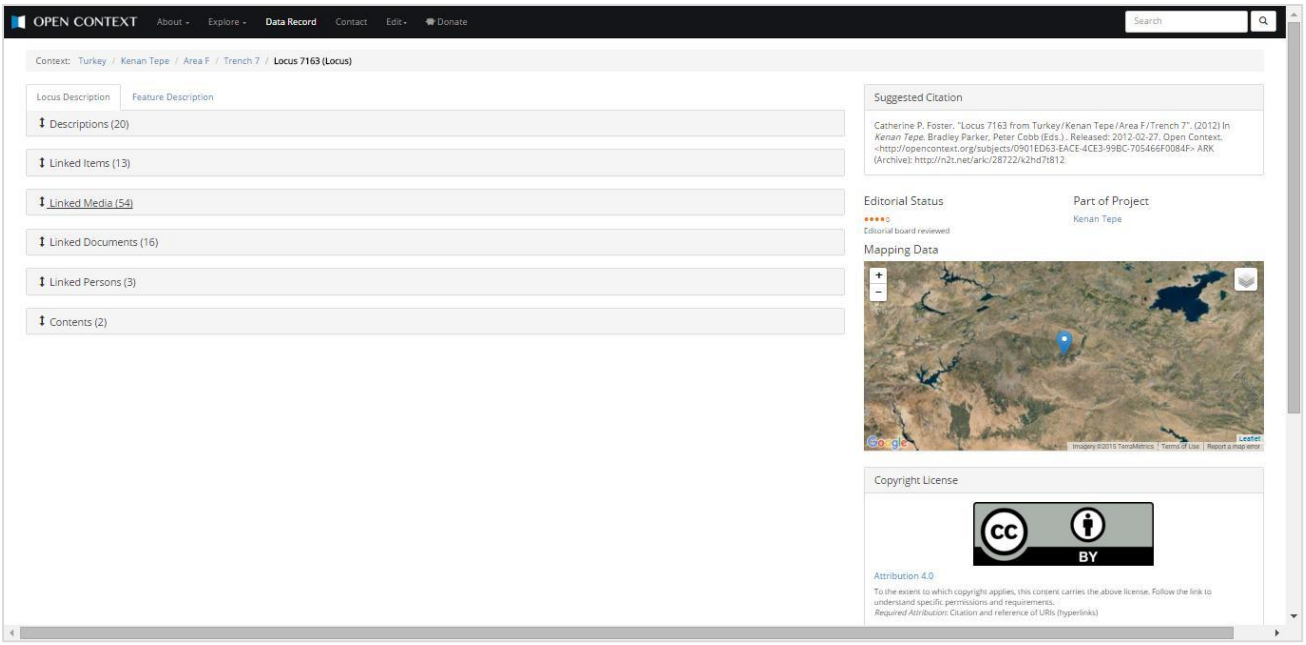
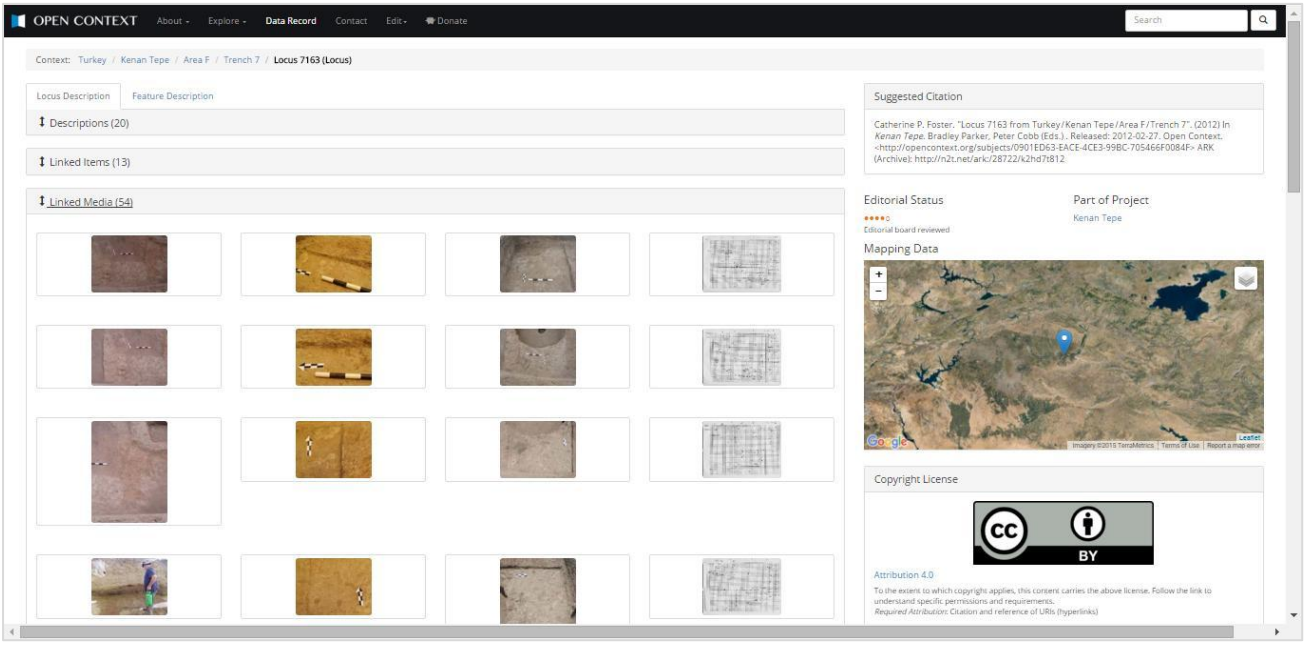



Figure 5c: In this view of the same item, the user has opened the "Linked Media" tab.