

Teaching Open Science: Published Data and Digital Literacy in Archaeology Classrooms

*Katherine Cook, Canan Çakırlar, Timothy Goddard, Robert Carl DeMuth,
and Joshua Wells*

Until recently, “reuse” in archaeological research most often referred to dusting off legacy collections in forgotten corners and basements, sifting through accumulations of gray literature, or applying new scientific techniques to test older interpretations. Advocates have long argued that this is critical to sustainable archaeological practice—that the pace of excavation, limited analysis, and underutilized but

data-rich collections create a missed opportunity for advancing knowledge and scholarship, not to mention being irresponsible and unethical (Bawaya 2007; Cook and Burchell 2012; Ferris and Cannon 2009). Despite ongoing pushes to safeguard collections in increasingly high-tech repositories and research facilities, and through discussions of best practice and changes to legislation, reuse of this kind remains

ABSTRACT

Digital literacy has been cited as one of the primary challenges to ensuring data reuse and increasing the value placed on open science. Incorporating published data into classrooms and training is at the core of tackling this issue. This article presents case studies in teaching with different published data platforms, in three different countries (the Netherlands, Canada, and the United States), to students at different levels and with differing skill levels. In outlining their approaches, successes, and failures in teaching with open data, it is argued that collaboration with data publishers is critical to improving data reuse and education. Moreover, increased opportunities for digital skills training and scaffolding across program curriculum are necessary for managing the learning curve and teaching students the values of open science.

El alfabetismo digital se ha citado como uno de los principales desafíos para la reutilización de datos y una mayor valoración de la ciencia abierta. Elemento clave para abordar esta cuestión es la incorporación de datos publicados en los programas formativos. En este artículo se presentan estudios de caso en el uso de plataformas de datos arqueológicos existentes en tres países (los Países Bajos, Canadá y Estados Unidos) para la enseñanza a estudiantes de diferentes niveles y habilidades. Al delinear planteamientos, éxitos y fracasos en la enseñanza con datos abiertos, se concluye que la colaboración con los editores de datos es fundamental para mejorar la reutilización de datos y la educación sobre los mismos. Además, es necesario aumentar las oportunidades de formación en habilidades digitales y el andamiaje a lo largo de los planes de estudios para administrar la curva de aprendizaje y enseñar el intercambio de datos y la reutilización como práctica arqueológica.

Advances in Archaeological Practice 6(2), 2018, pp. 144–156
Copyright 2018 © Society for American Archaeology
DOI:10.1017/aap.2018.5

limited (Ferris et al. 2015), most often relegated to small-scale graduate theses or preliminary research projects. Accessibility, quality of metadata, and institutional pressures likely influence these practices, but training opportunities, attitudes, and research structures also challenge those who reuse archaeological collections.

As in so many areas of archaeology, the internet has transformed reuse in theory and in practice. Publishing formats, copyright laws, access, preservation, and research ethics today are fundamentally shaped by electronic distribution methods, data management logistics, and even social media (Kansa 2012:498–499). So too have concepts of openness, comparability, and interpretation changed as the relationships between archaeologists and data evolve through digital practice (Lucas 2012:216; see also Shorish 2015:99). Seemingly the natural progression of earlier arguments for reuse, data sharing should remove the barriers that have limited the accessibility of existing collections to encourage further inquiry. However, despite predictions that primary data collection would give way to open data reuse (Beck and Neylon 2012:494; Naylor and Richards 2005:90), evidence of demand and impact remains limited (Beagrie and Houghton 2013; Huggett 2015:10). Atici and others have highlighted that,

while most see data sharing as an important goal, much attention focuses on problems relating to supplying researchers with data and less on how researchers can best consume and reuse data. . . . [F]ew studies have explored challenges researchers may face in the analysis of datasets produced by others [2013:664].

Even more concerning, Jeremy Huggett's musings on digital archaeology are both clear and cutting: "The fact remains that—whisper it quietly—few of us are actually using the archived digital data in the first place" (2016).

Digital literacy, in addition to data information literacy, is often cited as the primary hurdle to widespread adoption of digital data reuse in archaeological research (Shorish 2015:97–99). University classrooms, then, become a critical opportunity to ensure that future generations of archaeologists are comfortable with information management systems, thick/long data (Arbesman 2013; Wang 2013), and how to address the benefits and challenges of designing and managing digital data research projects. If we continue to teach archaeological practice in the same way, framed by traditional field and laboratory structures, we cannot expect skills, attitudes, and approaches to change in favor of reuse. This article brings together instructors and archaeologists involved in data publishing platforms to discuss experiences and barriers in using open data for teaching in three different countries: the Netherlands (Çakırlar), Canada (Cook), and the United States (Goddard et al.; Figure 1). We present experiences teaching at a range of levels, including undergraduate and graduate students, in advance of diverse career goals, including academic, government, and private-sector careers. These courses, which we developed independently, are used to develop collective

conclusions on the role of digital data in a range of archaeology training environments. In particular, this article considers the feasibility of teaching with published digital data, what students and instructors need from information management systems to ensure success in the classroom, and how to tackle the barriers that persist in education and their legacies beyond academic contexts.

OPEN ACCESS AND EDUCATION

In many ways, the movement for open access, defined as free of charge and free of many copyright and licensing restrictions, conceptualized by Suber (2012:4) as the barrier-free revolution, is perfectly suited to the demands of higher education. Discourse on open science parallels demands to remove economic and physical barriers to training within increasingly tense research and scholarly ecosystems (Kansa 2012:500; Suber 2012:29–39). In North America and parts of Europe, for instance, tuition increases alongside cuts to services, library resources, salaries, funding, and hiring rates have intensified economic strains in university research and teaching. Open-access data and publications reduce costs for high-quality content that would traditionally be transmitted via costly textbooks, library subscriptions, and hands-on or applied opportunities such as field schools and volunteer positions, which are not feasible for all students. As open-access resources are increasingly drawn upon to fill in these voids, students, particularly at the graduate level, find themselves responsible for advanced digital data and information management tasks that they may not feel prepared for. Archaeological training at all levels simply does not include enough data handling components. Open data are poised to transform the way we teach archaeology; however, inconsistent and limited training opportunities, typically as a stand-alone course or isolated workshop, continue to undermine the successes of open scholarship.

At the same time, the movement toward open data has been criticized for a lack of empirical evidence of demand or use in archaeology (Huggett 2015:10). Higher education could help to stimulate shifts in digital practice by providing the foundation for a more data-literate archaeological community that can move data management and reuse forward. Teaching with real case studies is always more relevant than using fake data, providing realistic experiences of the imperfections and complexities of archaeological evidence. Moreover, it is a chance to challenge future archaeologists to think differently about the process of archaeological research and reconsider traditional models of data production through field- and lab work. Finally, as Beagrie (2008) argues, increasing expectations for accessible and reusable data across disciplines and research streams, including commercial, government, and academic contexts, have precipitated a need for individuals trained in data management, preservation, and sharing. Teaching transferrable skills is already a focus for many archaeology programs today; however, advanced digital data skills are rarely included, despite their growing demand in many of the avenues for archaeology graduates.

Digital Literacy and Data Information Training

Digital literacy, then, is pivotal to promoting open data use, both in classrooms and beyond. At a recent forum at the Society for

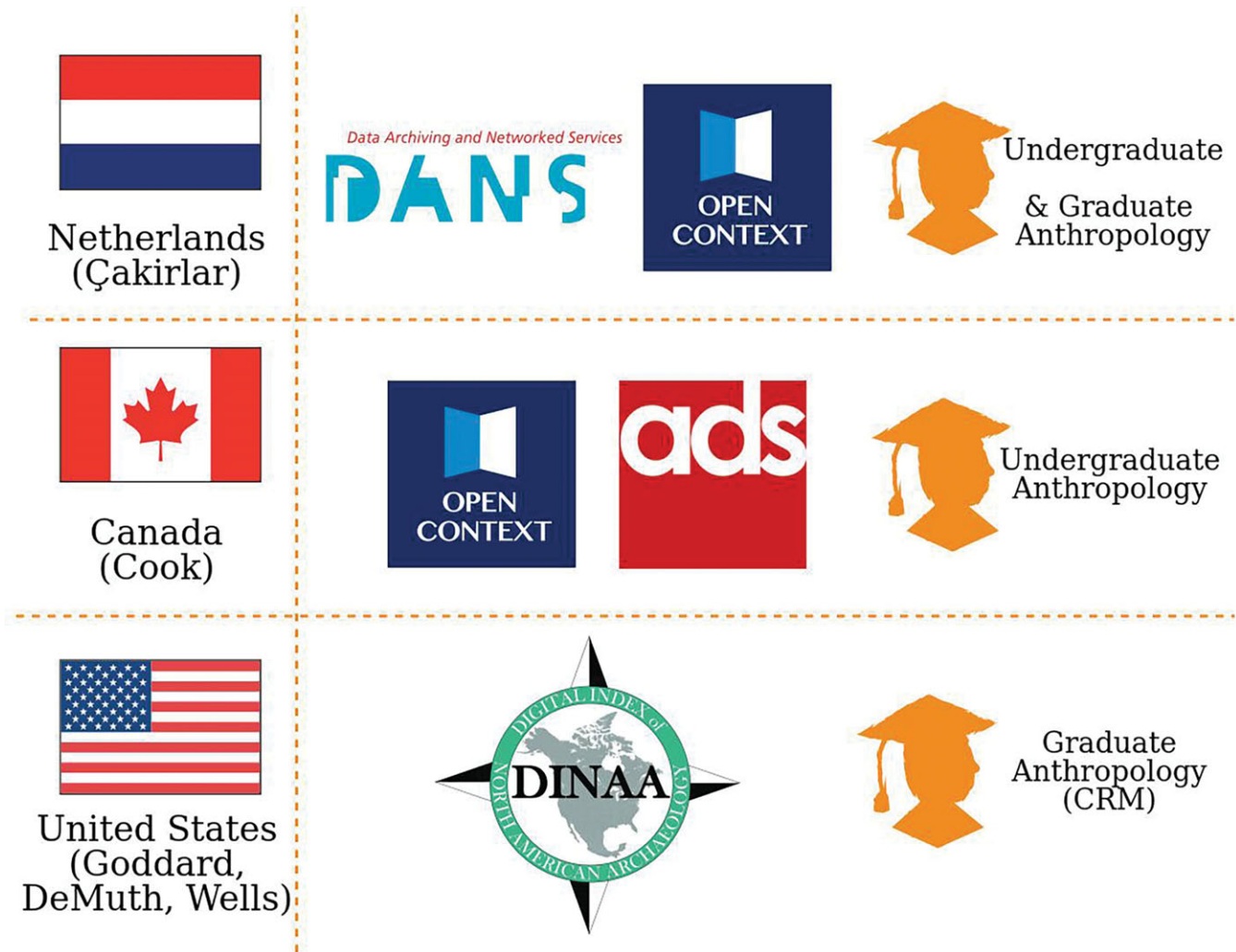


FIGURE 1. Summary of three case studies teaching with different datasets and students. ADS = Archaeology Data Service; CRM = cultural resource management.

American Archaeology meetings (Vancouver, 2017), “Beyond Data Management,” organized by Sarah Witcher Kansa and Eric Kansa, digital literacy was cited as one of the biggest barriers to data reuse by almost every panelist and many audience members. There are common expectations that computers will make data reuse easier, particularly for younger generations of archaeologists. However, even in the wake of the disintegrating “digital native” myth, information management training remains rare, while the increasing complexity of data, curation, and digital platforms challenges and limits reuse. As Naylor and Richards have observed,

Re-use of data requires a close understanding of the context of data collection and of the vocabulary used to describe the observations. The archaeologist of tomorrow needs training not so much in methods of data collection, but in data analysis and re-use [2005:90].

Carlson and others (2011) have identified a complex web of 12 competencies that are foundational to training effective digital researchers; these include not only skills in data processing, management, and preservation but also cultures of practice, ethics, and attribution (see also Carlson et al. 2013). Contextualized and critical discussions and applications of digital data production, curation, and reuse therefore must be incorporated progressively into undergraduate- and graduate-level classrooms to ensure not only training in the methods, software, and logistics of digital research but also informed considerations of theoretical opportunities and constraints.

Because of these challenges in resources, attitudes, and digital literacy, teaching with published data becomes a key context from which to consider the broader climate of open data in archaeology and the future of reuse and sustainable research practices. Promoting digital data reuse means new tools, skills, professional roles, and communication standards, transforming

most steps in the research process, from funding to workflow to citation (Kansa 2012:510). At the same time, the constant and rapid innovation of the Web environment challenges archaeologists to keep pace with active and thoughtful evolution of practice, infrastructure, and ethics policies (Kansa 2010:12). The following case studies aim to provide not only practical examples of the diverse methods of utilizing open data in classrooms but also broader considerations of the impact, constraints, and challenges both for higher education and for the open access movement.

OPEN ZOOARCHAEOLOGY AND BIOARCHAEOLOGY IN THE NETHERLANDS

The basic necessity to find meaning in large numbers of fragmented animal remains made zooarchaeology one of the most data-literate fields of archaeology early on (e.g., Clutton-Brock 1975; Uerpman 1978). The application of standards in recording techniques was proposed in the 1970s (e.g., Grant 1982; von den Driesch 1976) and widely adopted. Dutch zooarchaeology has been in the forefront of these developments, with Anneke Clason of the Groningen Institute of Archaeology highlighting the importance of distinguishing primary data from secondary data in zooarchaeology and urging zooarchaeologists to apply some general rules in publishing their observations (primary data) alongside their interpretations (secondary data [Clason 1972, which has since become a classic; see also Reitz and Wing 2008:153–250]).

In the age of digital revolution, it is crucial to teach how zooarchaeological data should be designed and managed for reuse. Zooarchaeologists worldwide have demonstrated the value of data sharing and reuse more vocally than ever before (Arbuckle et al. 2014; Atici et al. 2013; Kansa 2015). However, in the Dutch context, where the Dutch Science Foundation (the primary funder of Dutch science) requests that all research results be published open access, and all development-led archaeological research results have been deposited in an online repository (Data Archiving and Networked Services [DANS]) since 2007 in compliance with the Dutch Heritage Law, data sharing becomes imperative.

The Methodology

The availability of open-access zooarchaeological data on platforms such as DANS and opencontext.org facilitates the learning process. In the Department of Archaeology, University of Groningen, which offers zooarchaeology modules (courses) to undergraduate and graduate students, Çakırlar teaches the value of data reusability using methods in line with a hands-on, continuous learning approach. First, in lectures, syllabi, and reading lists, common data types collected, field-specific conventions and standards, must-haves of faunal reports (sensu Grigson 1978), and online resources for zooarchaeological data are introduced and discussed. Assigned readings, particularly for postgraduate students, include Atici and colleagues 2013 and Kansa 2015. While the lectures and readings are crucial to provide a formal framework and help students get familiar with some of the basic concepts, they lead to a passive learning atmosphere with little long-lasting effect.

TABLE 1. Class Exercise: Criteria for Peer-Reviewing Open Online Zooarchaeological Datasets.

#	Criterion
1.	Project description: Does the project description sufficiently describe the dataset?
2.	Organization/structure: Is the organization of the dataset logical? Coherent? Sensible? Does the dataset or project description have any major gaps or flaws?
3.	Longevity: Is there sufficient information provided with the dataset to ensure that the data can be understood by future users?
4.	Revision or referral: What are the most important changes that you judge to be necessary?
5.	Other comments: Is there anything else you like or do not like about the dataset?

Source: From S. Kansa, used with permission.

Second, students are asked to peer-review open online datasets and faunal reports “relevant to their research interests,” applying a set of criteria (Table 1) suggested by Sarah Whitcher Kansa (data scientist and zooarchaeologist) to the former and a set of criteria based on Historic England’s (2014) *Animal Bones and Archaeology. Guidelines for Best Practice* to the latter. Most students judge the datasets and reports they find as satisfactory (or pick the datasets and reports that they deem good enough in the first place). While they are “inspired” by this exercise, it does not help them to judge the reuse value of the published data, as they lack the incentive to truly reuse the data and because of the common tendency to accept published results as facts.

Third, students are assigned to extract and quantitatively explore open online zooarchaeological datasets and reports after having gained experience through analyzing clean datasets (e.g., performing simple quantifications such as comparing number of identified specimens results or constructing survivorship curves) presented by the lecturer and receiving points and feedback. The exercise has two aims: (1) to highlight that different students can reach different secondary data (e.g., minimum number of individuals) reusing the same set of primary data and (2) to aid answering a variety of research questions (e.g., What were the different animal husbandry strategies in Roman and non-Roman regions of the Netherlands? What were the main subsistence changes between late Paleolithic and Neolithic Turkey?). Now that the students are finally asked to reuse open, scholarly, “real,” sometimes peer-reviewed data (see Russell et al. 2013), they gain a fuller appreciation of what publishing should be like and what makes open data reusable. For a student, it may be difficult to imagine that different analysts may have massively different definitions of a zooarchaeological specimen, that even the concept of “specimen” can be interchangeable with “fragment” or “identifiable bone,” decreasing the compatibility of so-called basic quantification units such as the number of identified specimens across assemblages and analysts. By analyzing other researchers’ primary data, the importance of data compatibility and sufficient metadata becomes obvious, which is rare when working with faunal reports (secondary, interpreted data)—that is, of course, when the students are able to find open primary data online related to their research questions.

Challenges and Impact

When dealing with research questions pertaining to the Dutch setting, students search the online repository DANS, where data (a) are accessible after log-in, (b) are often embargoed by the authors, and (c) consist of secondary data (e.g., specialists' reports). The above-mentioned regulations do not distinguish primary from secondary data, and it is the norm in commercial archaeology for specialists to finalize their contact by submitting a report to project leaders. DANS has set some strict rules on how to organize the uploaded files but does not advise on their internal organization. Students often make do with the reports, rather than contacting the authors, which researchers at later stages with more time might do. When primary datasets are available, students face other challenges, such as missing metadata (e.g., What does *juvenile* mean as a descriptor of a femur fragment?), causing ambiguity; lack of standardization, for instance, due to typos; and lack of crucial data (measurements or standardized tooth aging). Frustrations of this sort ease when online primary data have been edited, linked, and/or peer-reviewed prior to publication (e.g., Open Context). Students are also challenged by this exercise due to their own lack of knowledge in information management systems (how to download a .csv file) or in statistical applications, which makes them feel intimidated with large differences in sample size.

While undergraduate and early postgraduate students say that they learn from dealing with good and bad datasets, later-stage postgraduate students who have to reuse other people's primary data are disappointed and worried that they will not reach their research goals. Undergraduate students find it difficult to appreciate that the present availability of online data is a major improvement from just a few years ago, when all data could still only be retrieved through personal connections and time-consuming archival and library research (if possible, as most zooarchaeological reports are in gray literature). Graduate students are "inspired to contribute to the online platforms" with their own datasets, and "seeing other people's primary data help[s] [them] look at online datasets and publications more critically" (from graduate students' course evaluations).

As a result of this program, throughout their zooarchaeology training students gradually learn the concept and the requirements of data reuse. They start appreciating the limitations of faunal reports and that producing reusable data, as Atici and others (2013) suggest, requires effort. However, most students remain illiterate in data management programs, and zooarchaeology courses alone cannot redress this deficiency. At Groningen University, the bachelor's curriculum in archaeology has recently been modified to include skills training necessary to prepare students to adopt the best practices in data management and data presentation in academic and professional settings, in the form of a compulsory module preceding the zooarchaeology modules. The new digital humanities master's program, in part designed by faculty members of the archaeology department, also offers modules suitable for archaeology students. A shift toward greater appreciation and more credit for reusable and reused data in both professional (heritage/commercial/salvage/contract archaeology) and academic/scientific settings will certainly encourage students and educators to place more emphasis on this crucial subject.

ADVANCED RESEARCH METHODS AND ANTHROPOLOGY IN CANADA

Considering increasing pressure on universities to offer courses providing hands-on research experience, many departments are creating opportunities for students to design and manage projects, foster independent and analytical thought, and develop transferrable skills such as time management, digital literacy, and leadership. The challenge for archaeology, as with many fields, is how to condense complex and long research processes, traditionally involving field- and/or laboratory work, into invariably short and intense semesters/terms for students. When it came to designing a new third-year research course for the Department of Anthropology at the University of Victoria (British Columbia, Canada), published data made it possible to access detailed and high-quality datasets that could be queried and analyzed within a semester, with the added benefit of building data information literacy and discussing open science in the process.¹

The Methodology

The course ANTH 319: Advanced Research Methods in Archaeology and Biological Anthropology was originally designed by Cook using Open Context, but in later versions she modified assignments to also use Archaeology Data Service to access more datasets, reflecting growing numbers of students and higher demands for data relevant to biological anthropology and bioarchaeology. The intensive research course introduced students to open data, how to access data in these particular interfaces, and how to make judgments about their research potential, before choosing their own dataset(s) to work with (Table 2). Students typically selected one to three datasets, often to compare sites, regions, or time periods to produce original interpretations. Research questions ranged from methodological and analytical in nature, for instance, evaluating the impact of the level of fragmentation of faunal remains on studies of butchery, to theoretical, for instance, exploring cremation as a reflection of differential personhood. After cleaning and coding the data, quantitative, qualitative, and statistical analysis was conducted. In the final stages, research results were presented in an academic-style paper and a piece written for general audiences.

Although students worked independently, they were placed in research groups that met weekly to discuss progress, challenges, and experiences, in addition to group discussions and one-on-one meetings with the instructor. This gave students the opportunity to see other types of datasets; examine the differences in quality, metadata, and organization; and reflect on implications for research. It also provided regular opportunities to check in on students' progress, experiences, and honest observations; address common problems; and consider the role of published data in the research process.

Challenges and Impact

Students' digital literacy was immediately and almost universally a challenge each time this course was offered. Despite streamlined and clean platforms designed by Open Context and Archaeology Data Service, with both basic and advanced search options, students quickly became lost in rabbit holes of projects, metadata, and arrays of file types. Technically advanced

TABLE 2. Workflow for University of Victoria ANTH 319 Research Projects.

Stage	Data-Driven Tasks	Feedback
Research design	Interface orientation/exploration	Informal with class
	Selection of dataset(s)	Informal with instructor
	Examination of metadata, published records/papers, etc.	Informal with small groups, instructor
	Design of original questions, methods	Formal: proposal assignment
Data preparation	Cleaning data (Open Refine tutorial)	Workshops with one-on-one support
	Coding, typologies and categorization	Informal research group
	Data management and organization	Formal: dataset submission
Data analysis and interpretation	Data analysis and statistics software	Workshops with one-on-one support
	Data interpretation	Informal research group, instructor
		Formal: data analysis assignment
Communication	Data presentation	Workshops with one-on-one support
	Writing with data	Informal research group, instructor
		Formal: written paper(s) assignment

concepts such as linked data, application programming interfaces, GeoJSON data, and so on quickly overwhelmed students. These elements and data formats, however, are critical to what has been termed “Mash-Up-Ready” Data—that is, content that can easily be used in different applications or combined with other sources (Kansa 2010:12; Kansa, Elliott, et al. 2010) to pursue original research questions. With a lack of resources written in plain language or designed for students/beginners, there was a significant investment in delivering this content in class and during office hours, not to mention developing course resources and workbooks, to explain the technology, interfaces, and terminology as well as their relevance to anthropology and how to make the most of them in research. This quickly led to frustration, anxiety, and confusion that temporarily derailed progress for many students and took the focus away from the primary course goals of independent research design and management. This stage can easily alienate researchers, whether students or professionals, particularly if they are not confident in their digital literacy skills and problem solving. It quickly became clear in teaching this course that further focus is needed on communicating how to use digital frameworks, which, although designed to make data more useful and accessible to researchers, are exceedingly complex.

Digital literacy, however, did not in fact turn out to be the most difficult barrier to selling open data to student researchers. Expectations of what anthropological research looks like in addition to attitudes regarding worth, research impact, and contributions were routinely in conflict with the values of open science, data reuse, sustainability, and ethics that were foundational in this course’s design (Figure 2). Although some may argue that the attitudes of undergraduate students do not necessarily reflect the attitudes of the professional discipline, open discussions in class did point to more deep-rooted problems than simply popular culture representations of research.

There were few examples in the literature of high-impact research using published data and none that the upper-level students were familiar with prior to the course (see Arbuckle et al. 2014; Bevan 2012; Kintigh 2006; Snow et al. 2006). Moreover, few of the

studies that are based on the aggregation and interpretation of large-scale, open datasets include detailed discussion of their research methods, particularly choices in cleaning, unifying, and analyzing data to address issues of comparability and compatibility. Even more concerning was the pervasive impression that funding, for both students and professionals alike, is more likely to support research involving original field- or lab work than the analysis of existing data. Although systematic evidence was not used to support this argument, the universal belief among students that funding, employment, publication, professional recognition, and ultimately success would be undermined by pursuing research with published data does speak to wider issues in how we recognize, value, and measure the impact and place of this type of research within archaeology, not to mention communicate the realities of archaeological research.

Some of the barriers experienced in teaching this course can be addressed within university curriculum. Greater scaffolding to introduce published data, digital concepts and skills training, and diverse research processes more gradually and methodically throughout numerous courses within programs is critical to the success of students. Currently work is under way to incorporate a smaller, more guided project using Open Context datasets in a second-year, required course at the University of Victoria. It is hoped that this will make students more familiar with not only the technical side of accessing digital data but also the values and ideas behind sharing data, earlier in the anthropology curriculum. Research utilizing published data should also be highlighted in introductory archaeology courses, where students are first given a taste of what the archaeology research process looks like (often limited to excavation, laboratory, and occasionally archival research). Lecture and seminar discussions can be bolstered through reading lists that include reuse case studies for thematic and advanced skills courses.

However, some of the barriers extend beyond what can be designed into curriculum. Resources for teaching that make technical jargon more accessible, for instance, are essential for these types of courses to become more commonplace. The challenge of passing on courses like ANTH 319 to other instructors,

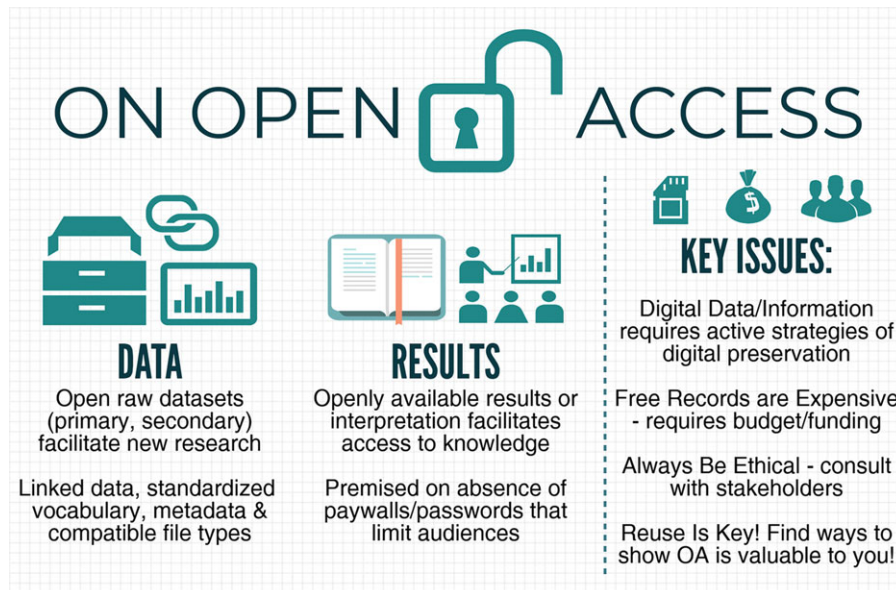


FIGURE 2. Detail of a workbook from ANTH 319, outlining open access in anthropology.

with less experience with digital archaeology, highlights the need for further collaboration between instructors and data publishers, including training opportunities. Finally, the expansion of open-access textbooks and educational resources complements the ethos of openness, increases the potential to produce high-quality media communicating open data and digital archaeology practice and theory to students, and is, indeed, imperative.

CULTURAL RESOURCE MANAGEMENT AND PUBLISHED DATA IN THE UNITED STATES

The Digital Index of North American Archaeology (DINAA) is a project to develop a comprehensive hub for archaeological site data across North America (Wells et al. 2014). DINAA integrates these differentially sourced data into an open interoperable dataset via a controlled, common vocabulary. To account for site security, DINAA assigns each site to a raster grid system and removes all sensitive information from the data (including site locations) published on Open Context (Wells et al. 2014). In this way, DINAA creates an open, publicly available tool that can be used to examine archaeological site data from the macro perspective without putting any legally protected data at risk.

In an expansion of public outreach, the DINAA team (Wells and DeMuth) partnered with an instructor (Goddard) at Adams State University (ASU) to produce educational teaching aids for ASU's (2017a) online cultural resources management (CRM) master's program. Adams State started this master's program in 2013 to serve the critical need of CRM professionals seeking continuing education to further their careers who did not have traditional options available to them. The initial teaching aids, discussed in this case study, were developed for use in CRM 510: Technologies and Techniques. CRM 510 is a survey course focused

on exposing students to "the wide variety of technologies and the necessary techniques to implement those technologies in the field and/or office in a CRM setting" (ASU 2017b). The primary foci of the course are the use of technology in "the field, analysis, and data dissemination" (ASU 2017b). The course is divided up into weekly modules on ASU's Blackboard page for CRM 510. Each week, students are provided with a series of readings, activities, and online video lectures/presentations and are required to respond to discussion questions in the classroom forums.

The Methodology

Two stand-alone geographic information system (GIS) tutorials and two educational modules were developed using DINAA's vocabularies and informational ~20-km raster grid system, which does not use or contain site coordinates.² The GIS tutorials are integrated with a lesson on digital data analysis and were designed as a basic introduction to the DINAA spatial interface. Students are asked to navigate to a specific data query in DINAA and download a GeoJSON file using a supplied URL. The goal of these exercises is to demonstrate how DINAA data may be used and how to access the data. One tutorial was made to introduce how to use DINAA data in QGIS, and the other was designed for Google Earth.

The two educational modules were developed to complement the CRM 510 lesson on digital data dissemination. These lessons give a more thorough introduction to DINAA through tutorials requiring students to explore the graphical user interface menus to access descriptive information within the DINAA dataset. Both modules are hosted on Indiana University's course management system, Canvas (Figure 3). The first module introduces DINAA, serving as a basic user guide to the DINAA graphical user interface. The primary assignment in the first module requires students to research site forms from two states repre-

2016

View All Pages

Published

Edit

Settings

Home

Announcements

Assignments

Discussions

Grades

People

Pages

Files

Syllabus

Outcomes

Quizzes

Modules

Conferences

Collaborations

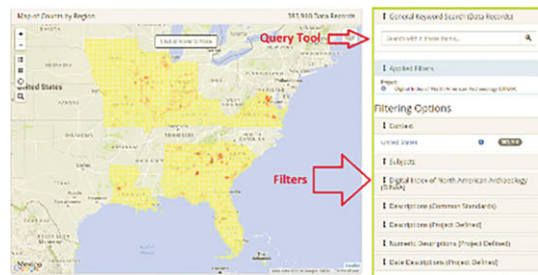
Chat

Library
Resources

Using DINAA

Now that you are comfortable with the Tools, its time to start using DINAA's querying and filtering tools.

These are both located to the right of the map area as shown below.



First you should explore the filters, we'll use this more in the next module but basically there are 2 kinds of basic filters in the DINAA project - Common Standards and Project Defined. Common standards are the parts of the database that DINAA built to make multiple state site file systems work together. Currently the primary common standard are the Descriptions. This consists of the 'translating' system that ties individual state culture-history terms to our modified version of the CIDOC-CRM ontology.

FIGURE 3. View of a lesson in the first module created for Adams State University CRM 510.

sented in DINAA and then examine how those forms influenced the ways in which each state's data are represented in DINAA.

The second module involves advanced exploration and use of the DINAA dataset. Students are asked to perform specific queries, as well as access and explore individual data records. The goal is to demonstrate how the DINAA dataset can be of use to archaeologists. The module culminates in a scavenger hunt in which the students are asked to perform specific tasks using DINAA and answer questions from their resulting data.

Challenges and Impact

Working with DINAA and the modules provided gives real-world direction and examples of some of the bigger issues faced by CRM practitioners. The ASU students cover a broad range of CRM experience, from recent college graduate to more than 20 years in the profession. The overall topic during this portion of the class is about data dissemination. DINAA is a specific example of a more general discussion of websites and databases. Often in the course students ultimately focused on the data being presented and lost sight of the general technology and methodology for dissemination. Although DINAA is not designed for CRM use, it is an educational resource built in

part with information from CRM activities, so the students saw the ultimate power and usefulness of this type of tool, as well as potential challenges (both technically and politically) to making such a system work.

Students indicated that despite some minor glitches in the interfaces they all recognized the power of archaeological data synthesis as a practice. The discussions in class also explored other types of Web databases, including other forms of spatial WebGIS-type sites, as Goddard has particular experience in this area. As they relate to their own field of study, the students were eager to work with government agencies to develop more useful tools like DINAA for their respective states to contribute to DINAA directly. It is clear that working with published data put the successes and challenges of this work in the minds of students, and some have continued to pursue solutions to moving this type of technology forward. Several students were inspired to seek out more experience in archaeological data management through their own research or in internships to work on the idea of digital data dissemination in various forms.

Working with ASU provided several benefits for the DINAA project, because the CRM 510 students had a diverse set of backgrounds, experiences, and plans for their archaeological careers. They were able to provide important feedback regarding how DINAA worked well and how it could be improved (Figure 4).

CRM 501 Student Forum Feedback

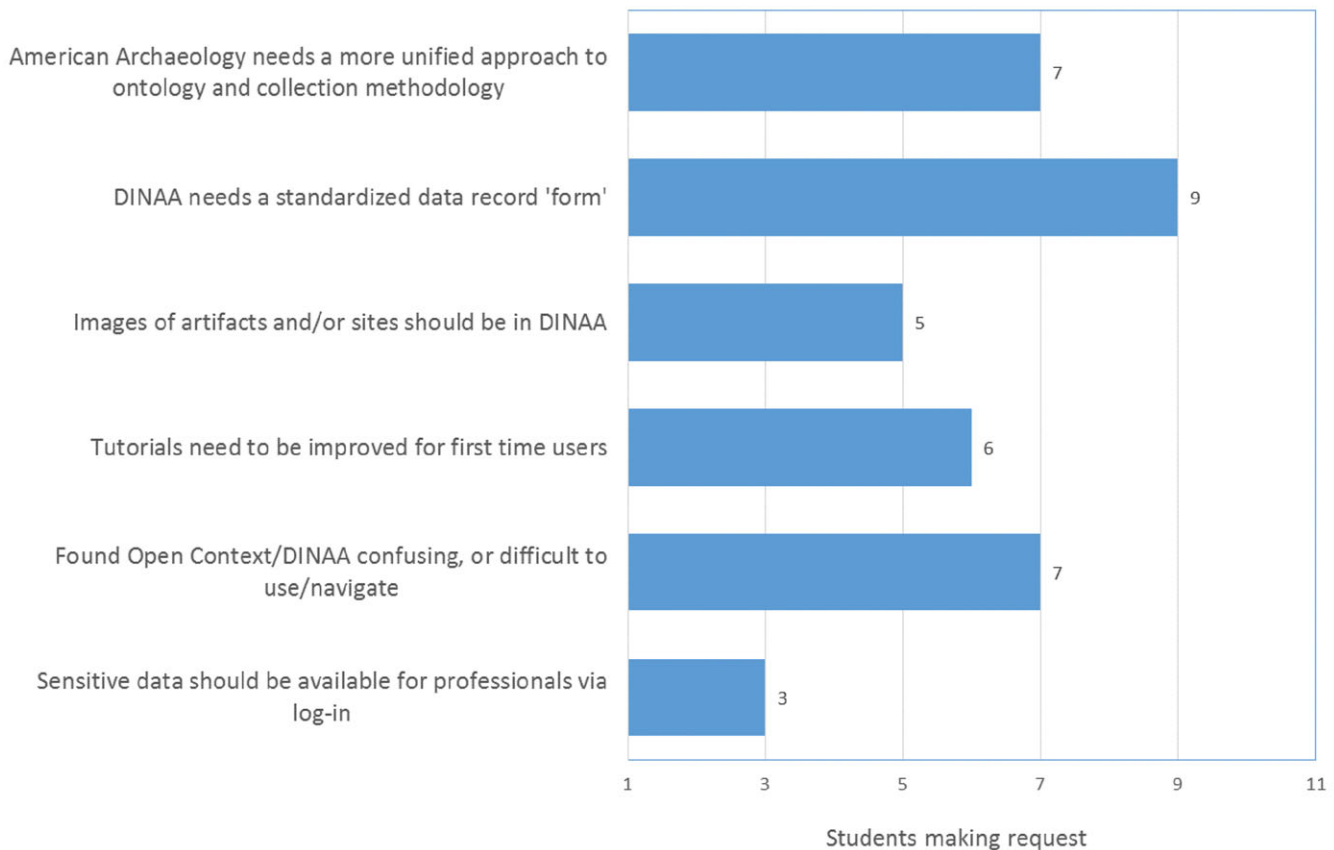


FIGURE 4. Linear bar graph of Adams State University student feedback on the Digital Index of North American Archaeology (DINAA).

For example, ASU students were quick to grasp the importance of DINAA as an example of comprehensive vocabulary and why such a tool is now necessary. One student stated: “Another thing that I would like to see improved is the vast inconsistencies in terms. Every project innovating their own terminology makes extracting data from DINAA a nightmare.”³ Additionally, another student later stated:

The best part about the database is the partnerships it is creating. The states have a central hub to relay site data. . . . The public has a look into the government’s way of collecting data. Researchers can connect to sites around the country and the world.

ASU students also provided constructive feedback. For example, currently when a user selects a site file in DINAA the site essentially provides elements of the basic data entered into a state site reporting form. Because each State Historic Preservation Office manages its own form, there is little immediate standardization of these data. Many students found the lack of standardization in DINAA’s presentation of such data confusing. They further debated how useful a unified site data record within DINAA would be, recognizing that while such an undertaking would be

helpful, it would be difficult to implement. Such considerations of design complexity mentioned by ASU students will help designers consider future growth of the DINAA system. The students also noted functionality issues more broadly oriented to access, such as lags in query completion and the rendering of maps (Figure 5); such issues involve combinations of query complexity and Web latency, which can be considered as DINAA matures. Some students found the user guide and tutorial confusing for a first-time user and made helpful editorial suggestions. A few found the entire system confusing—those in this subgroup often self-disclosed low computer literacy. Such comments suggest that DINAA may serve as a tool in the long term to promote archaeological computer literacy but that more specialized development is needed.

DISCUSSION

Just as analysis of how researchers use published primary data (Atici et al. 2013) is critical to designing better platforms and publishing better data, collaborations between data publishers, instructors, and students help to address barriers to data reuse for future generations of archaeologists while also developing understandings relevant to increasing sustainability and impact

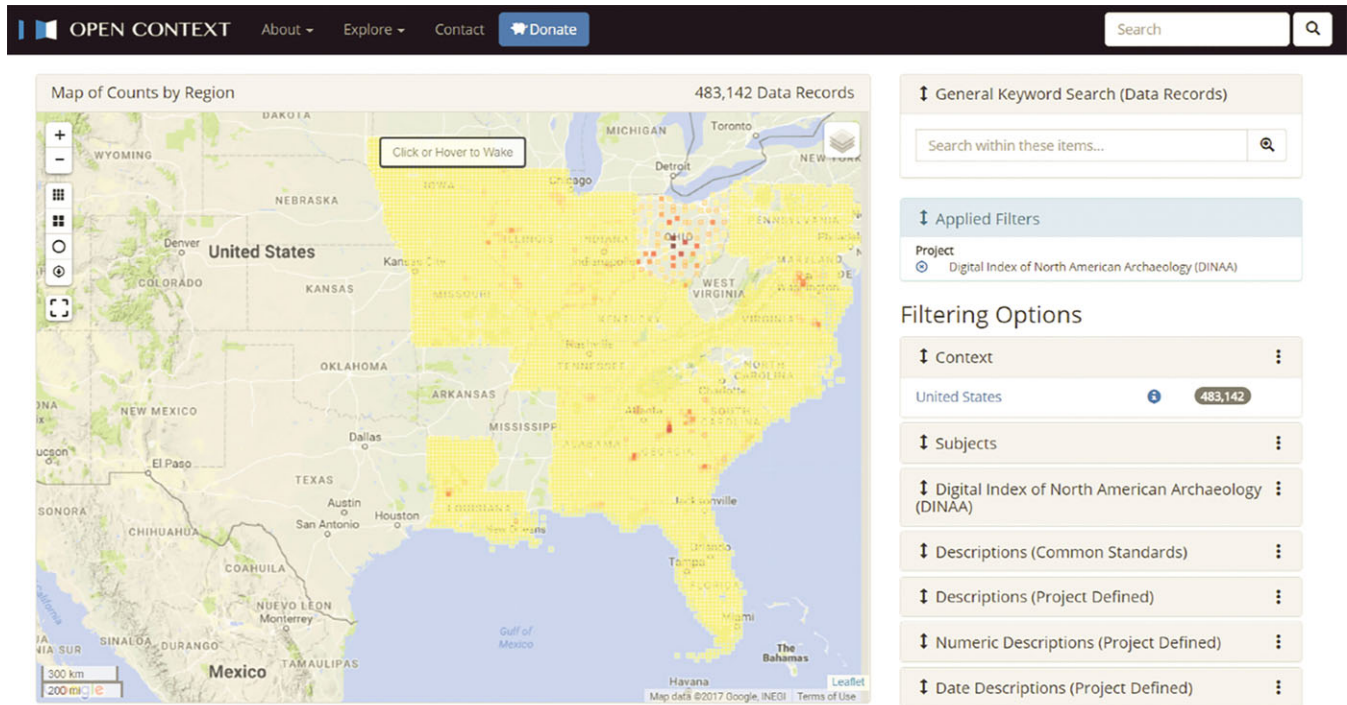


FIGURE 5. Overview of the Digital Index of North American Archaeology map within Open Context.

within the discipline more broadly (Figure 6). For open data to be incorporated into classrooms more seamlessly, there is a need for more high-quality, easily digestible resources. The recentness of methodological and theoretical attention to open data in archaeology (Huggett 2015:14; Kansa, Kansa, et al. 2010:303), and the resulting lack of consistency and agreement, has no doubt contributed to the limited availability of resources for teaching. Detailed examples of digital data-based research are starting to emerge, providing valuable and reproducible examples (see, for instance, Marwick 2017). However, the case studies above also found collaboration among instructors, students, and data publishers to create resources that correlate with specific interfaces invaluable in making open data accessible for students at all levels. DINAA (2017), for instance, has open-access information sheets for using the digital index for teaching and for students, in addition to the open modules described above. These are concise information sheets, easy to download and print, designed to help introduce open data and how to navigate specific platforms, to integrate into courses at a range of levels.⁴ Open Context has also collaborated with Shawn Graham (Carleton University) to encourage student innovations in data visualization through the Open Context and Carleton Prize for Archaeological Visualization. This competition challenged individuals and teams, with special student categories, to create reproducible methods and applications, with the view to raising awareness of open data. Other solutions to these issues are under development, for instance, the ODATE: Open Digital Archaeology Textbook Environment (Shawn Graham, Neha Gupta, Michael Carter, and Beth Compton) is currently being built, with sections on making data useful and methods for cleaning data, using linked open data, and data publishing (Graham 2017; Graham et al. 2018). As an online environment, ODATE proposes to create an e-textbook accompanied by a digital laboratory to engage students in what

Graham defines as the essence of digital archaeology: “the creative use of primarily open-source and/or open-access materials to archive, reuse, visualize, analyze and communicate archaeological data” (2017). Finding creative solutions to removing barriers (economic, technological, etc.) for students to access high-quality resources must be at the core of open archaeology development strategies.

Attitudinal shifts are also critical in these contexts. Kansa warns,

The discipline should not continue to tolerate the personal, self-aggrandizing appropriation of cultural heritage that comes with data hoarding. . . . [F]ailure to incentivize greater data transparency would demonstrate an egregious failure of leadership and utter dysfunction in a discipline supposedly devoted toward building and preserving knowledge of the past [2012:507; see also Carson 1996:316; Huggett 2015:9].

The differences in practices, requirements, and stewardship of archaeological data in the Netherlands, Canada, and the United States impact the context and concepts of best practice embedded in curriculum and students’ attitudes. The Dutch Science Foundation’s requirements for published data and the complex relationship of tiered government agencies, CRM, and academic archaeologists in North America not only shape the types of open data available and their quality and accessibility but also filter into students’ outlook on data reuse. Although progress has been made in dismantling the ivory tower, there is still a long way to go toward recognizing that classrooms and students are not isolated but, rather, deeply entangled in professional and public attitudes toward open science and the legislation and policy that



FIGURE 6. Recommendations for archaeologists involved in higher education, data publishing, and research for encouraging data reuse and literacy.

structure data publishing. These case studies highlight not only the impact of these contexts on the quality and quantity of open data but also the ways that they shape reuse within and beyond academia. Evidently, far more critical analysis of these structures needs to be pursued to evaluate and inform best practice in data publishing and reuse. More extensive and systematic consultation with and feedback from students and instructors is critical in moving forward.

In the history of the discipline, few inventions have so challenged archaeologists, including instructors and researchers alike, to keep pace in the rapid evolution of research, training, and publishing methods and theory as the internet has. The values of sustainability and reuse have long been threads in archaeological practice and thought, but now, more than ever, we need critical

and creative solutions. Although technology has advanced to provide the opportunities for the openness and critical reengagement with archaeological data, our own frameworks for research and training have not yet caught up. Advanced digital literacy is critical to fully benefit from innovative information management systems designed to create more compatible data, detailed metadata accompaniments, and intricate systems of linked and geolocated data. The alignment of training programs with these recent technological developments can only be achieved with (1) systems of curriculum scaffolding for progressive skills development; (2) the production of high-quality, routinely updated resources and case studies detailing methods, theory, and results; (3) training for instructors and professionals; and (4) enhanced communication and collaboration between diverse avenues and roles pursued by archaeologists today. Further challenges remain

for information management platforms in how best to support, track, and measure this form of demand and reuse, given that classroom use may not result in expected forms of citations and metrics but does have a high level of impact in shaping ideas and practice. As a complex ecosystem, archaeology is a tangled web of researchers, instructors, students, and now, more than ever, information management teams, governments, and community stakeholders; our strength will come from collectively challenging, engaging, integrating, and collaborating on building foundational skills, method, and theory for the future of reuse in archaeology.

Acknowledgments

The development of the Digital Index of North American Archaeology is funded by the National Science Foundation (Award #1217240, 1623621, 1216810, 1623644) and the Institute of Museum and Library Services (Award #LG-70-16-0056-16).

Data Availability Statement

The Digital Index of North American Archaeology online course modules are available at <https://iu.instructure.com/courses/1591811>. Materials for the open data courses offered at the University of Victoria are available at <https://github.com/KatherineRCook/TeachingOpenData> (DOI:10.5281/zenodo.1009169). Other data referenced in this article are available through the Open Context, Archaeology Data Service, and Data Archiving and Networked Services publishing platforms.

REFERENCES CITED

- Adams State University
 2017a MA in Humanities with a Cultural Resource Management Emphasis. Electronic document, <https://www.adams.edu/academics/hgp/cultural-resource-management/>, accessed June 15, 2017.
 2017b Cultural Resource Management Emphasis. Electronic document, http://catalog.adams.edu/contentm/blueprints/blueprint_display.php?bp_listing_id=154&dont_show_done=1&blueprint_id=160&sid=23, accessed June 15, 2017.
- Arbesman, Samuel
 2013 Stop Hying Big Data and Start Paying Attention to “Long Data.” *Wired Magazine* 1(21) <https://www.wired.com/2013/01/forget-big-data-think-long-data/>.
- Arbuckle, Benjamin S., Sarah Whitcher Kansa, Eric Kansa, David Orton, Canan Çakırlar, Lionel Gourichon, Levent Atici, Alfred Galik, Arkadiusz Marciniak, Jacqui Mulville, Hijkje Buitenhuis, Denise Carruthers, Bea De Cupere, Arzu Demiregri, Sheelagh Frame, Daniel Helmer, Louise Martin, Joris Peters, Nadja Pöllath, Kamilla Pawłowska, Nerissa Russell, Katheryn Twiss, and Doris Würtenberger
 2014 Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals across Neolithic Turkey. *PLOS ONE* 9(6): e99845 <https://doi.org/10.1371/journal.pone.0099845>.
- Atici, Levent, Sarah Whitcher Kansa, Justin Lev-Tov, and Eric C. Kansa
 2013 Other People’s Data: A Demonstration of the Imperative of Publishing Primary Data. *Journal of Archaeological Method and Theory* 20(4):663–681.
- Bawaya, Michael
 2007 Archaeology: Curation in Crisis. *Science* 24(5841):1025–1026.
- Beagrie, Neil
 2008 Digital Curation for Science, Digital Libraries, and Individuals. *International Journal of Digital Curation* 1:3–16.
- Beagrie, Neil, and John Houghton
 2013 The Value and Impact of the Archaeology Data Service: A Study and Methods for Enhancing Sustainability. Archaeology Data Service. Electronic document, <http://archaeologydataservice.ac.uk/research/impact.xhtml>.
- Beck, Anthony, and Cameron Neylon
 2012 A Vision for Open Archaeology. *World Archaeology* 44(4):479–497.
- Bevan, Andrew
 2012 Spatial Methods for Analysing Large-Scale Artefact Inventories. *Antiquity: A Quarterly Review of Archaeology* 86(332):492–506.
- Carlson, Jake, Michael Fosmire, Chris Miller, and Megan R. Sapp Nelson
 2011 Determining Data Information Literacy Needs: A Study of Students and Research Faculty. *Portal: Libraries and the Academy* 11: 629–657.
- Carlson, Jake, Lisa Johnston, Brian Westra, and Mason Nichols
 2013 Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project. *International Journal of Digital Curation* 8(1):204–217.
- Carson, Cindy Alberts
 1996 Laser Bones: Copyright Issues Raised by the Use of Information Technology in Archaeology. *Harvard Journal of Law and Technology* 10(2):281–319.
- Clason, Anneke T.
 1972 Some Remarks on the Use and Presentation of Archaeozoological Data. *Helinium* 12(2):139–153.
- Clutton-Brock, Juliet
 1975 A System for the Retrieval of Data Relating to Animal Remains from Archaeological Sites. In *Archaeozoological Studies: Papers of the Archaeozoological Conference 1974, held at the Biologisch-Archaeologisch Instituut of the State University of Groningen*, edited by Anneke T. Clason, pp. 21–34. North-Holland, Amsterdam.
- Cook, Katherine, and Meghan Burchell
 2012 Envisioning New Approaches to Archaeological Practices in Ontario. *OAS Arch Notes* 17(6):5–8.
- Digital Index of North American Archaeology
 2017 DINAA Information Sheets. Open Context. Electronic document, <http://ux.opencontext.org/archaeology-site-data/>.
- Ferris, Neal, Rhonda Bathurst, Michael Carter, and Namir Ahmed
 2015 Sustaining Ontario’s Archaeology Digitally. *Heritage Matters* 13(2):8–11.
- Ferris, Neal, and Aubrey Cannon
 2009 Capacities for a Sustainable Archaeology. Paper presented at the Ontario Archaeological Society Annual Meeting, Waterloo, Ontario, Canada, October 17.
- Graham, Shawn
 2017 ODATE: Open Digital Archaeology Textbook Environment (Original Proposal). *Electric Archaeology* (blog), February 6. Electronic document, <https://electricarchaeology.ca/2017/02/06/odate-open-digital-archaeology-textbook-environment-original-proposal/>.
- Graham, Shawn, Neha Gupta, Michael Carter, and Beth Compton
 2018 *The Open Digital Archaeology Textbook Environment*. Draft version. Electronic document, <https://o-date.github.io/draft/book/index.html>.
- Grant, Ann
 1982 The Use of Tooth Wear as a Guide to the Age of Domestic Ungulates. In *Ageing and Sexing Animal Bones from Archaeological Sites*, edited by B. Wilson, Caroline Grigson, and Sebastian Payne, pp. 91–107. BAR British Series 109. BAR, Oxford.
- Grigson, Caroline
 1978 Towards a Blueprint for Animal Bone Reports in Archaeology. In *Research Problems in Zooarchaeology* 3, edited by Don R. Brothwell, Kenneth T. Thomas, and Juliet Clutton-Brock, pp. 121–128. Institute of Archaeology, London.
- Historic England
 2014 Animal Bones and Archaeology: Guidelines for Best Practice. Electronic document, <https://content.historicengland.org.uk/images-books/publications/animal-bones-and-archaeology/animal-bones-and-archaeology.pdf>

- Huggett, Jeremy
2015 Digital Haystacks: Open Data and the Transformation of Archaeological Knowledge. In *Open Source Archaeology: Ethics and Practice*, edited by Andrew T. Wilson and Ben Edwards, pp. 6–29. De Gruyter Open, Berlin.
- 2016 Digital Data Realities. *Introspective Digital Archaeology* (blog), June 29. Electronic document, <https://introspectivedigitalarchaeology.wordpress.com/2016/06/29/digital-data-realities/>.
- Kansa, Eric C.
2010 Open Context in Context: Cyberinfrastructure and Distributed Approaches to Publish and Preserve Archaeological Data. *SAA Archaeological Record* 10(5):12–16.
- 2012 Openness and Archaeology's Information Ecosystem. *World Archaeology* 44(4):498–520.
- Kansa, Eric C., Tom Elliott, Sebastian Heath, and Sean Gillies
2010 Atom Feeds and Incremental Semantic Annotation of Archaeological Collections. In *Computer Applications and Quantitative Methods in Archaeology—CAA 2010*, edited by J. Melero and P. Cano. CAA, Fargo, North Dakota.
- Kansa, Eric, Sarah Witcher Kansa, M. M. Burton, and C. Stankowski
2010 Googling the Grey: Open Data, Web Services and Semantics. *Archaeologies* 6:301–326.
- Kansa, Sarah Witcher
2015 Using Linked Open Data to Improve Data Reuse in Zooarchaeology. *Ethnobiology Letters* 6(2):224–231.
- Kintigh, Keith W.
2006 The Promise and Challenge of Archaeological Data Integration. *American Antiquity* 71:567–578.
- Lucas, Gavin
2012 *Understanding the Archaeological Record*. Cambridge University Press, Cambridge.
- Marwick, Ben
2017 Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory* 24:424–450.
- Naylor, J., and Julian D. Richards
2005 Third-Party Data for First Class Research. *Archeologia e Calcolatori* 16:83–91.
- Reitz, Elizabeth, and Elizabeth S. Wing
2008 *Zooarchaeology. Cambridge Manual in Archaeology*. Cambridge University Press, Cambridge.
- Russell, Nerissa, Katheryn Twiss, Hylke Buitenhuis, Sheelagh Frame, Louise Martin, Çatalhöyük Zooarchaeology, Lisa Yeomans, Arek Marciniak, Kamilla Pawłowska, Claire Christensen, David Orton, Benjamin S. Arbuckle, Arzu Demirergi, Stephanie Meese, Levent Atici, Stuart Campbell, Elizabeth Carter, Liz Henton, Banu Aydinuloglu, Alfred Galik, Adam Watson, Denise Carruthers, Ian Cameron, Rhian Mayon-White, Amanda Erwin, and Vesna Dimitrijevic
2013 EOL Computational Data Challenge: Primary Zooarchaeology Dataset. In *Çatalhöyük Zooarchaeology, Ilıpınar Zooarchaeology, Çatalhöyük Area TP Zooarchaeology, Zooarchaeology of Neolithic Ulucak, Domuztepe Excavations, Erbaba Höyük and Suberde Zooarchaeology, Köşk Höyük Faunal Data, Zooarchaeology of Öküzini Cave, Zooarchaeology of Karain Cave B, Barçın Höyük Zooarchaeology, Pınarbaşı 1994: Animal Bones, Çukuriçi Höyük Zooarchaeology*, edited by David Orton, Nerissa Russell, Katheryn Twiss, Louise Martin, Sheelagh Frame, Hylke Buitenhuis, Arek Marciniak, Benjamin S. Arbuckle, Levent Atici, Canan Çakırlar, Stuart Campbell, Elizabeth Carter, Alfred Galik, and Denise Carruthers. Open Context. Electronic document, <http://opencontext.org/tables/f07bce4fb08cfe926505c9e534d89a09>. DOI:10.6078/M75H7D6D.
- Shorish, Yasmeen
2015 Data Information Literacy and Undergraduates: A Critical Competency. Information Literacy beyond the Library. *College and Undergraduate Libraries* 22(1):97–106.
- Snow, Dean R., Mark Gahegan, C. Lee Giles, Kenneth G. Hirth, George R. Milner, Prasenjit Mitra, and James Z. Wang
2006 Cybertools and Archaeology. *Science* 311(5763):958–959.
- Suber, Peter
2012 *Open Access*. MIT Press, Cambridge, Massachusetts.
- Uerpmann, Hans-Peter
1978 The “Knocod” System for Processing Data on Animal Bones on Archaeological Sites. In *Approaches to Faunal Analysis in the Middle East*, edited by Richard H. Meadow and Melinda A. Zeder, pp. 149–167. Peabody Museum Bulletins 2. Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge, Massachusetts.
- von den Driesch, Angela
1976 *A Guide to the Measurement of Animal Bones from Archaeological Sites*. Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge, Massachusetts.
- Wang, Tricia
2013 Big Data Needs Thick Data. *Ethnography Matters*, May 13. Electronic document, <http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/>, accessed January 18, 2018.
- Wells, Joshua J., Eric C. Kansa, Sarah W. Kansa, Stephen J. Yerka, David G. Anderson, Thaddeus G. Bissett, Kelsey Noack Myers, and R. Carl DeMuth
2014 Web-Based Discovery and Integration of Archaeological Historic Properties Inventory Data: The Digital Index of North American Archaeology (DINAA). *Literary and Linguistic Computing* 29(3):349–360. DOI:10.1093/lc/fqu028.

NOTES

1. Teaching materials, including the syllabus and assignments, for this course can be accessed at <https://github.com/KatherineRCook/TeachingOpenData> (DOI:10.5281/zenodo.1009169).
2. The online course modules featuring DINAA are publicly available; those wishing to utilize them in their own courses can access them at <https://iu.instructure.com/courses/1591811>.
3. In fall 2016 the Adams State Institutional Review Board approved the use of student feedback as long as students were informed of their right to choose to participate of their own volition in the follow-up questions about their experience with these modules and that their grade was not affiliated with their feedback. Students were also informed that their names would not be used in the presentation of their comments and data. In some cases students did not submit their feedback until after the semester was completed.
4. Other open data repositories, such as Open Context, Archaeology Data Service, and tDAR: The Digital Archaeological Record, also have user help guides and FAQs. The format and length of these guides, however, are heavily tied to the ease with which they can be integrated into classroom settings.

AUTHORS INFORMATION

Katherine Cook ■ Department of Anthropology, University of Victoria, 3800 Finery Road, Victoria, British Columbia, V8P 5C2, Canada (katherinecook@uvic.ca, corresponding author)

Canan Çakırlar ■ Groningen Institute of Archaeology, Groningen University, Poststraat 6, NL-9712 ER, Groningen, the Netherlands

Timothy Goddard ■ Department of History, Anthropology, Political Science and Philosophy, Adams State University, 208 Edgemont Boulevard, Alamosa, CO 81101, USA

Robert Carl DeMuth ■ College of Arts and Sciences, Indiana University Bloomington, 107 South Indiana Avenue, Bloomington, IN 47405, USA

Joshua Wells ■ Department of Sociology and Anthropology, Indiana University South Bend, 1700 Mishawaka Avenue, South Bend, IN 46615, USA