

Digital Archaeological Data Collection and Archaeological Data Repositories:

How Digital Environments are leading to a Convergence of Data Collection Standards

Julia Silver

Jsilve37@jhu.edu

(410) 660-5009

Johns Hopkins University

Museum Studies Digital Curation Certificate Program

December 5, 2015

Abstract

This paper discusses how the lack of standardized methods for on-site data collection has impacted the ability of archaeologists to preserve archaeological data and make it accessible for reuse. Repositories and databases, like the Alexandria Archive Institute's (AAI) Open Context and Digital Antiquity's tDAR (the Digital Archaeological Record), have gained traction as prominent platforms to both publish archaeological data and to store and make accessible raw data. However, gaps remain between data collection and data accessibility, which seem unlikely to be closed by the adoption of standards for the documentation and description of field data. The following questions are considered: What are the challenges to standardization in the collection and documentation of field data? In the absence of agreed-upon standards, what alternative methods can aid in the preservation and accessibility goals of archaeologists? One promising development, a mobile platform created by the Federated Archaeological Information Management Systems (FAIMS) Project, is considered as an alternative path that may lead to a convergence of data collection standards in the future.

Table of Contents

- I. Introduction
- II. Research Methodologies
- III. Historical Background
 - a. Defining Digital Archaeology
 - b. DigMaster
 - c. The Beginning of Online Data Publication
 - d. The Archaeological Information Ecosystem
- IV. The Lack of Standardization Regarding On-Site Data Collection
 - a. What Accounts for the Lack of Standardization?
 - b. The Impact of a Lack of Standardization
 - c. Alexandria Archive Institute's Open Context, Data Publication, and the Archaeological Data Lifecycle
 - d. Digital Antiquity's tDAR
- V. The FAIMS Project
 - a. Scope and Purpose
 - b. Challenges in Creating and Implementing the Mobile Platform
 - c. The Potential of the FAIMS Mobile Platform
- VI. Conclusion: The Future of Digital Archaeology
- VII. Resources

*Digital Archaeological Data Collection and Archaeological Data Repositories:**How Digital Environments are leading to a Convergence of Data Collection Standards***Introduction**

Like many other disciplines in the cultural and science sectors, archaeology is advancing towards a completely digital model. However, with the exception of electronic and digital tools used in field excavation, progress has been slow. One reason the field has been slow to adopt a comprehensive digital model is the lack of a standardized method of on-site data collection. There is such diversity in the field of archaeology that archaeologists have had to new methodologies to meet the specific needs of their sites for each individual project. The lack of standardized methods of data collection lends itself to other challenges that have an impact on the field. The result is that much archaeological data, both digitized and born-digital, has been inaccessible to other researchers. Inaccessibility of archaeological data also hinders the preservation of the data that is created and recorded for future use. The varied needs of archaeologists and the uniqueness of each site have prevented the development of a single standardized model. The variety of specific needs of archaeologists and their projects is an inherent deterrent to the reuse of digital archaeological data and the preservation of that data beyond what is in direct support of a published field report.

Within the last three years, an alternative solution has arisen and gained traction that does not impose standard methods of data collection on archaeologists as a prerequisite to digital data collection. The Federated Archaeological Information Systems (FAIMS) project [<https://www.fedarch.org/>] has been developing “discrete, federated mobile and web applications for the creation, refinement, archiving, and dissemination of digital data” (Ross, 2015, p.112). Within the FAIMS ecosystem the

possibility exists for semantically interoperable datasets to arise through the convergence of data collection and the federation of data to existing archaeological data repositories.

This paper presents the FAIMS model as a potential solution to the absence of a standard method of on-site data collection. Through partnering with existing archaeological repositories like Digital Antiquity's The Digital Archaeological Record (tDAR) [<https://www.tdar.org/>] and the Alexandria Archive Institute's (AAI) Open Context [<http://www.opencontext.org/>], a digital archaeological ecosystem is being created wherein archaeological data can be seamlessly transmitted from the site to an open access database. A digital archaeological ecosystem is the goal of organizations like FAIMS and AAI. In the words of Eric Kansa and Sarah Kansa:

"Many, many different individuals and organizations create and manage archaeological data. Data must be able to flow across institutional boundaries in order to be synthesized and understood as a whole. The research community therefore needs to consider the broader *information ecosystem* that extends beyond institutional and disciplinary boundaries, and encompasses multiple information systems" (Kansa & Kansa, 2014, p.226).

As Jacobs & Holland (2007) speculated, "That reality about archaeological work—in the field, the lab, and the study—makes it imperative that archaeologists, state archivists, antiquities authorities, and museum curators share data openly and freely. To do so is responsible investment in the future of the discipline" (p. 198). More recently, Kansa & Kansa (2013) further discuss the implications of a lack of data professionalism in regards to data collection and dissemination in their article "We All Know That a 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communications." They point out, "In archaeology, a discipline that relies upon destructive research methods, lack of information sharing not only inhibits scholarship, but also represents a tragic loss of irreplaceable historical knowledge" (p.1). Therefore, digital data collection on-site will be proposed as a simple pathway to creating high

quality field data that can be readily transferrable to open-access repositories for the purpose of archiving and reuse; which are two important components of the digital data lifecycle that will enable the progression of the field of archaeology as a whole.

Research Methodologies

This paper researches the origin of the challenges that have faced the archaeological community in regards to adopting a single standard method of on-site data collection and explores the possibilities offered by a digital data collection platform, such as FAIMS has developed. First, context is given to the term “digital archaeology” as it relates to this paper. Then, a historical background of the field of digital archaeology is presented using a brief review of published literature and a meta-analysis of the progression of the field of digital archaeology to how that is relevant today. Finally, interviews were conducted with a number of experts in the field: Dr. Jodi Reeves Flores, the Senior Digital Curator of tDAR at Digital Antiquity; Dr. Sarah Witcher Kansa, the Executive Director of The Alexandria Archive Institute, which created and operates Open Context; and Dr. Adela Sobotkova, the Development Coordinator of the FAIMS Project. The interviews are used to support the analysis of relevant literature and to acquire a deeper understanding of the current status and direction of the field of digital archaeology, a field in which the landscape is changing on a daily basis, and where even current published literature can become outdated very quickly. They will also add a personal component from those currently working with the challenges and experiencing the successes their organizations have encountered. Additionally, the interviews will provide an insight into the possible direction in which the field may be headed within the very near future.

Historical Background

Defining Digital Archaeology

To begin a discussion on digital archaeology, one must first define what exactly digital archaeology entails. Ashely Richter's (2014) article "So What is Digital Archaeology?" explains the various meanings that the phrase can assume. She states that,

"[Digital Archaeology] is a bridge between the archaeological experts who want analytical tools and the wider public who (we hope) want to be engaged in the past. It is an interdisciplinary space where computer science and cultural heritage blend interchangeably as they work towards a common goal."

For the purpose of this paper, the term "digital archaeology" goes beyond implementing the use of digital tools in the field to collect data. It also includes the curation of all archaeological data: both data that is digitized as well as born-digital data. The act of curating data simply involves understanding the basic workflow of how archaeological data is produced, processed, stored, and made accessible. Subsequently, whenever the term "data" is referenced, it refers to digital archaeological data and includes data that is both digitized and born-digital.

Another definition of digital archaeology¹ is presented in *Archaeology 2.0: New Approaches to Communication and Collaboration*, edited by Eric Kansa, et al (2011). In the volume they discuss how digital archaeology and Archaeology 2.0 relates to the concept of Web 2.0:

¹ An alternative definition of "digital archaeology" discusses "Personal Digital Archaeology," a term defined by The Digital Archaeology Institute (n.d.). Personal Digital Archaeology is a "framework that seeks to help people capture, preserve and share their life stories, through sustainable learning material. It also helps people rescue digital artifacts from obsolete forms." For the purpose of this paper, this is not a definition that will be considered.

“Virtual (or digital) archaeology is a powerful tool for visualizing and understanding archaeological data as well as for producing and communicating it to the public (Evans and Daly 2006: 253). It is also an educational resource for the general public and students in archaeology and heritage management. Many re-creations of greatly detailed archaeological sites have been created with standard modeling, rendering, and animation techniques. Digital archaeology makes possible increased rates of publication of archaeological materials through the use of the Internet. Its “open-source knowledge” allows researchers to quickly and inexpensively produce and communicate archaeological knowledge to a broad community of international specialists, schools, and the interested public alike and even to get them interactively involved in this process” (p.125).

This definition is relatable to Ashely Richter’s and incorporates the importance of publishing archaeological data outside of just the excavation report. Additionally, the definition also involves the idea that archaeological data should be “open-source knowledge” and is important in the FAIMS Project section of this paper.

DigMaster

Archaeologists have historically utilized new technology as they have become available for use on-site during archaeological data creation and collection. However, it was not until the late 1990’s that the notion of making the data accessible to the general public and other researchers, who were perhaps not archaeologists and did not have the proper training to interpret the data in the same way, began to surface. One such example of archaeological data being presented online in an open-source platform is DigMaster, developed by Paul Jacobs and Chris Holland. In their article “Archaeology Online: New Life for Old Dead Things” (1997) they explained that

“DigMaster is an online experiment in the publication of archaeological materials. The intent is to test whether electronic publishing will meet several needs of the archaeological community, including a more robust presentation of archaeological data, prompt publication, and collaboration between researchers working on related excavation projects.”

The various successes of DigMaster are outlined in the article. The successes of using a digital platform to publish an entire archaeological artifact collection and site stratigraphy include: the timely manner in which the project was completed, and the fact that the project occurred in a fiscally-sound fashion. Jacobs goes on to state that “although [at the time] none of the three figurine sets has yet reached final publication (and may not for several years), the artifacts have already served the needs of others working in the field.” The significance of what DigMaster accomplished involved the ability to collaborate with other archaeologists and researchers who could use the published data to help interpret their own findings. Jacobs was also able to show how useful online publication could be and how, although there is an initial investment of time and resources, the end result not only supported his archaeological project but those of other archaeologists as well. The DigMaster project marks the evolution of project-based publication of datasets to entire online platforms dedicated to digital data publication.

The Beginning of Online Data Publication

The beginning of archaeological data publication and data curation with the creation of open-access databases like DigMaster, which is still serving the archaeological community today, paved the way for online data publication in the field of archaeology as a whole. Because of the first initiatives and the need to make collections accessible to others working in the field for the purpose of on-site collaboration (with the added benefit of transparency and reuse by researchers and the public), organizations today like AAI have been able to publish and make accessible over a million records in

Open Context alone (Alexandria Archive Institute, n.d.). AAI is a California-based non-profit organization that created and operates Open Context, "...a digital library-backed system developed by AAI for Web-based publication of research data" (Alexandria Archive Institute, n.d.) and archives data for preservation purposes with the University of California, California Digital Library.

Digital archaeology, however, does not begin nor end with online data publication. Online data publication can occur more efficiently and can be the result of utilizing digital methods to record archaeological data on-site. For instance, if the archaeological data is collected using a digital program, the transition to publication can occur, in some cases, at the push of a button. The digital archaeology workflow begins in the field, and can occur during the first on-site survey before digging even begins. Ixchel Faniel, et al (2013) discuss the role of data collection procedures during data reuse in "The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse." Faniel, et al studied the impact that initial data collection had on the context of that data. The case study included: how the data was interpreted in the field, how data collection impacted the context that was destroyed as a normal result of excavation, and how different approaches archaeologists employ impact the overall context of the data (p. 7). The overall consensus was that,

"Having access to data collection procedures helped respondents understand and verify the data against the archaeologists' research objectives and interpretations. The respondents also relied on archaeologists' presentation of the documents created during field work, the reputations of the archaeologists, their scholarly affiliation, and the institutions where the data were housed for additional insight into the data" (p.7).

The impact that a lack of proper context can have in data interpretation have begun to be addressed in recent years as digital platforms are developed that have the capability to record site activity and excavation at every step. Digital data collection platforms leave the guesswork out of what context the

collection methods have on the data and simply present the raw data as it is discovered. The context includes the methodologies and ontologies the archaeologist employed during excavation and the main research questions the archaeologists set out to answer.

The Archaeological Information Ecosystem

In addition to the adoption of online data publication, which has the added benefit of allowing archaeologists complete control over how their raw data is edited and published, the idea of an archaeological information ecosystem began shortly after the creation of the first online archaeological databases. Sarah Kansa and Eric Kansa (2014) explain the archaeological information ecosystem in the article “Data Publishing and Archaeological Information Ecosystem”:

“To make data an integral aspect of twenty-first century archaeological understanding, we need to take a wider view of the lifecycle of research data in the larger information ecosystem. To simplify a complex reality, stages in a data lifecycle include: planning, collection, analysis, and publication” (p.224).

The archaeological information ecosystem includes each stage of the data lifecycle and the digital platforms that support each of the stages.

Keith Kintigh (2006) discusses in his article “The Promise and Challenge of Archaeological Data Integration” the results of a National Science Foundation-funded workshop “that focused on the integration and preservation of digital databases and other structures data derives from archaeological contexts” (p.567). Kintigh discusses what would need to occur to create a “cyberinfrastructure” for archaeological data. The benefits of such a system are the same as those echoed today and include data access and reuse, integration, and preservation (p.572-573). The vision of the goals that Kintigh describes is “An effective information infrastructure for archaeology would encourage the research use

of existing data, satisfy the data requirements of integrative and synthetic research, facilitate entry of data into the infrastructure, and provide for the preservation of irreplaceable data” (p.573). Kintigh also outlined the challenges of a cyberinfrastructure. The challenges consist of the initial development of the technology involved and its maintenance, having adequate resources to create and maintain the infrastructure, and the adoption of the program by archaeologists and researchers (p.575-576). Furthermore the challenges mirror those that are apparent today in regards to the implementation and integration of programs, software, and databases that exist for the various parts of the archaeological data lifecycle.

In the next two sections, Kintigh’s challenges of a cyberinfrastructure will continue to be discussed as they are still present today. Although digital data publication has gained traction in the field, using entirely digital methods for initial data collection has progressed at a slower pace. The digital methods for data collection discussed in this paper refer to entirely digital data collection workflows rather than just the occasional use of a digital implement. The reasons for the slower-pace adoption of comprehensive, reusable digital data collection platforms are similar to the hurdles that data publication platforms and repositories have faced, despite the expansive possibilities of accessibility, reuse, and collaboration that exist when they are utilized. The initial concept of what is included in the archaeological information ecosystem is being expanded every day and yet the same issues that have resulted in a lack of a standard method of data collection are still present.

The Lack of Standardization Regarding On-Site Field Data Collection

What Accounts for the Lack of Standardization?

Methodologies in archaeology vary greatly depending on the region, time period studied, the type of site or survey, and a multitude of other factors. The differences cause archaeologists to create a set of standards specific to their sites that will work for them and are applicable to the initial interpretation of the data that is generated. Outside of the fact that every archaeological site is unique, the reasons why archaeologists utilize different methods are numerous and include: governing regulations specific to the country in which they are digging; different requirements depending upon the needs of the universities or organizations that support and fund the projects; the use of varied metadata schema that may 'fit' the descriptive and administrative needs of the site more fully; and even simply, the questions the archaeologists are asking in regards to their site can affect how detailed their documentation and reports need to be. Thus, varied methodologies lead to different concepts and data schemata, which hinder comparison and may require tedious (manual) conceptual mapping before data reuse is possible.

However, it is not merely a lack of standard methods of data collection that poses a road block to reuse and preservation. The lack of standard methods of data collection is functionally related to the lack of standardized ontologies. In the context of archaeology, ontology refers to the representation of archaeological knowledge. According to Faniel, et al (2013), the focus on ontology development in archaeology "centers on classification...and semantic inference" (p.10). Similar to how a single standard method for on-site data collection is not the solution to the ultimate goal of the preservation and reusability of archaeological data, neither are more fully-developed ontologies. Faniel, et al (2013) explains that their results indicate the obstacle to data reusability resides in the complex and error-prone field procedures. Faniel states,

"...ontologies, even if more fully developed to describe the conditions of data creation can only address part of archaeology's data management challenges. Our results illustrated complex

records management and inventory control needs, as artifacts and samples move through several steps from the ground to museum storage facilities. Each step introduces new chances for mishaps that may alienate the artifacts from or garble key contextual information. As archaeologists increasingly come to use shared datasets, there will be increasing need for better managed and reliable data creation processes throughout the entire research cycle” (p.11).

Therefore, the consensus here is not about creating new standards, methodologies, and ontologies. It is about how the process with which raw archaeological data is created will come to dictate the reusability of the data, how well the data is preserved for future archaeologists and researchers, and even how the data can be linked to the artifacts uncovered at a site. Linked data is especially important since the artifacts from archaeological digs typically reside in a museum, separate from the original contextual data.

The Impact of a Lack of Standardization

Levent Atici, et al (2012) discuss the need for better initial documentation of all archaeological data (both paper and digital) in their article “Other People’s Data: A Demonstration of the Imperative of Publishing Primary Data” in which independent researchers study a zooarchaeological legacy dataset. For researchers to reuse and accurately interpret legacy datasets left by other archaeologists, adequate recording of the data that is generated needs to occur first to “improve confidence of future reuse of data” (p.21). The study revealed that because of the lack of original documentation regarding the context of the data, (wherein the context refers to: what original research questions were asked, and what methodologies and ontologies were employed), that their researchers felt uncertain about asking questions, especially those directly related to the context of the data (p.21). From the study, a link was created between the quality of initial documentation of archaeological data and the ability to further interpret, reuse, and even preserve the data. Varied approaches to data collection do not just have an

impact on the difficulty of preserving and interpreting that data. They also have a direct effect on the ability of archaeologists and researchers to reuse that data. Raw data itself, without the proper context, contains very little meaning.

Alexandria Archive Institute's Open Context, Data Publication, and the Archaeological Data Lifecycle

Given the lack of standards in regards to all aspects of field data collection that directly impacts the context of data, the various organizations that operate open access archaeological repositories each take slightly different approaches to making archaeological datasets accessible. Eric Kansa, et al (2013) conducted a case study of data integration and reuse that is documented in their article "Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology." The result of the study concluded with editing the raw archaeological data, wherein the researchers each took the amount of time they needed to prepare their datasets for dissemination (p.62). The researchers are not alone to edit, annotate, and interpret their data before it was published to Open Context. Open Context has editors that assist archaeologists with reviewing and cleaning datasets. Part of the mission of Open Context, according to their website (n.d.) is to, "publish your data and preserve it with leading digital libraries. We take steps beyond archiving to richly annotate and integrate your analyses, maps and media. This links your data to the wider world and broadens the impact of your ideas." What also occurs is the preservation of some of the context associated with the original, raw data that is collected. As Sarah Kansa says, "Our work aims to not only ensure that research content is preserved, but that it is done so in a way that makes it more likely to be understood and reused in the future" (personal communication, October 22, 2015).

The online publication of archaeological data is becoming a part of the lifecycle of the data generated by fieldwork and a large part of the archaeological information ecosystem. While publishing data online is a viable solution for many in the field whose goal is reuse and preservation, there are still

inherent challenges outside of the time and resources it takes to prepare raw data for publication. One of the primary challenges comes back to initial data collection. As Kansa & Kansa (2014) discuss:

“If researchers lack incentives to disseminate data, they will invest less thought and effort in how to create and manage their data at the start of their investigations. Data they collect will tend to be messier, more error-ridden, and more difficult to interpret. Similarly, if their colleagues do not regularly disseminate data, ideas about good data organization, modeling and management practices will not be shared. Without sharing experiences about what works, we run the risk of having researchers continually “reinventing the wheel,” repeatedly making costly and unnecessary mistakes” (p.224-225).

The challenge of adopting newer data collection methods will be discussed further in the next section. However, Sarah Kansa mentioned “The data we’re publishing will only realize its full potential when it can participate in a thriving ecosystem of similar projects. Otherwise, it’s like having the first telephone – who are you going to call?” (personal communication, October 22, 2015). It is also important to note that an issue with implementing digital data collection methods entails ensuring the quality of the data that is collected. In regards to traditional methods of on-site data collection (i.e. using paper forms or a simple notepad and pencil) much can be lost, misinterpreted, or unknowingly duplicated from the time it takes the data to leave the field and be entered into even the simplest of research databases. By assisting archaeologists and researchers with cleaning up and editing their data before it is published, Open Context can circumvent some of the challenges with handling a variety of data that was collected using a variety of approaches.

Digital Antiquity’s tDAR

Another approach to archaeological data reuse is Digital Antiquity's tDAR platform. Digital Antiquity is located at Arizona State University and is supported by the Arizona State University Libraries and the Arizona State University Digital Repository for the purpose of long-term preservation and accessibility. With less administrative guidance from Digital Antiquity than AAI, no formal review system that results in publication, and a user-friendly interface, tDAR allows archaeologists and researchers to upload their own data and datasets, whether they have edited them or not. According to its website (2013), tDAR functions "as a repository for digital data" and "The core mission of tDAR is to support better stewardship of the documents and data produced during research." The tDAR repository allows archaeologists and researchers to use tDAR to manage and organize their data and download reports. There is less focus on the quality of the data and the context of the data within the larger scope of any one particular site or the field. However, tDAR "is dedicated to the teaching of archaeology." This method for collecting, storing, and reusing archaeological data is simpler in nature but still includes the challenges regarding the quality of the raw data that is collected and the initial method of data collection. Jodi Flores explains that "The...challenge here is keeping tDAR general enough for all areas and sub-disciplines, while making the tool and the metadata we collect still useful" (personal communication, October 8, 2015). tDAR has likewise found the approach to keeping the platform general to be successful, as there are currently 8,000 registered users. Although they are not sure how many people are actively using the publically-accessible data, tDAR's staff knows that a vast majority of their user base are those who wish to simply access the data or reuse it rather than contribute to it. It is important to note that accessibility and reuse is only made possible by the preservation feature of the database (Jodi Flores, personal communication, October 15, 2015).

Although there are a number of other digital repositories, online databases, and digital publishing platforms for archaeological data (such as: KORA, The Mukurtu Project, The OCHRE Data

Service, HEURIST, etc.) the focus on tDAR and Open Context in this paper is due to the differing approaches each takes. These two unique repositories have proven to be successful tools for both archaeologists and researchers seeking to reuse the archaeological data and store archaeological data collected from a digital platform. However, the lack of a standard method of data collection still poses a challenge, especially in regards to the quality and reusability of the data. In the next section, the use of the FAIMS model will be discussed as a possible solution to the loss of quality and context of raw archaeological data once it leaves the hands of the archaeologists who generated the data and finds its way to an existing open-access repository like the two mentioned.

The FAIMS Project

Scope and Purpose

A new possible solution to the lack of standardization in data collection, and the absence of even standardized archaeological ontologies and methodologies, has emerged over the last few years. The FAIMS website (2015) explains that they “build tools for digital data collection in the field, and online processing and archiving of the resulting data.” Of specific interest to digital data collection on-site is the FAIMS Mobile Platform, which is a “scalable and flexible solution that can be relied on for all types of digital fieldwork.” The FAIMS Mobile Platform is a free app download available in the Google Play Store from any Android mobile device. In addition to allowing data to be created and recorded using their platform, the software itself is open source and the individual modules available for use within the app are also open source. The FAIMS digital tools only focuses on data collection and the data and datasets are federated to existing repositories, like tDAR and shortly, Open Context. The federation of data allows the data to flow easily from one platform to another.

As Shawn Ross, et al (2015) discusses in "Building the Bazaar: Enhancing Archaeological Field Recording Through an Open Source Approach," there was not any software that existed for mobile platforms that met the needs of the FAIMS' stakeholders (who include Australian and overseas archaeologists), thus "The mobile data collection platform was the only component that we decided to build from scratch..." (p.112). The FAIMS Project conducted extensive surveying and held a workshop to determine what requirements would need to be met with a mobile platform and what challenges they would face. The end result is an "archaeology-specific tool for the collection of well-structured digital data in the field and laboratory" and "The mobile platform is flexible enough to accommodate archaeologists' idiosyncratic needs and practices" (p. 113).

Challenges in Creating and Implementing the Mobile Platform

Adela Sobotkova discussed in an interview the three main challenges the FAIMS Project faced in their initial stages of development. The challenges included solving technical problems, financial restrictions, and adoption (personal communication, November 10, 2015). The primary technical problem was the creation of a platform that would actually be used. It was discovered through consultation with potential users that a generic, static data logger would not accommodate the field as a whole and would wind up unutilized. So the highly customizable data collection platform that "should allow for the collection of spatial, structured data and multimedia offline, have hooks for linked data and semantic interoperability, could always be traced to its author, and could never be lost...[while being] fairly easily editable" was the end result (personal communication, November 5, 2015). The second challenge was to finance the initial development of the Mobile Platform and to keep the project financed. Sobotkova explained that "What we did was pool the resources of 42 institutions (who committed in-kind time) and apply for matching government funding," a process that had to be repeated two years later (personal communication, November 5, 2015). "FAIMS has been led by the

University of New South Wales, Sydney, in collaboration with participants from 40 organizations, including universities, archaeological consultancies, and heritage agencies in Australia and overseas” (Ross & et al, 2015) and has since relocated to Macquarie University in Sydney, Australia.

The third challenge is one that requires a shift in the way that archaeologists work in the field, and involves the adoption of a digital data collection method like the FAIMS Mobile Platform. The mobile app has been successfully used and open source data collection modules now exist within the app, making it easier for certain projects to more easily switch from their current method of data collection to the FAIMS Mobile Platform. However, it still requires archaeologists to dedicate time and effort upfront to switch from a paper-based system to a digital one, and it can be a major transition to suddenly have a lot of hardware for a previously paper-based project (personal communication, A. Sobotkova, November 10, 2015).

The cost of hardware is the only cost that falls solely on the shoulders of program directors; the costs associated with deployment support are absorbed by FAIMS. Additionally, digital forms are not flexible like paper forms; directors need to have data structures well worked out ahead of time, which involves a lot of pre-field work (personal communication, November 10, 2015). However, the adoption hurdle can be met once the benefits of switching to the FAIMS Mobile Platform are experienced first-hand and the platform is recommended by colleagues. In addition, more published case studies and reports of the platform’s success in the field in widely read journals will aid in further adoption of the mobile platform.

The Potential of the FAIMS Mobile Platform

One of the main benefits of the FAIMS Mobile Platform that sets it apart from other digital data collection methods is that all of the software is open source. Archaeologists, once choosing to

implement the Mobile Platform into their field research, can create modules to record the various aspects of their site, or do some light editing to the existing modules, if one already suits their needs. Additionally, the modules ensure that all of the fieldworkers who are recording the data do so in exactly the same fashion, and instead of having to ask a site director for guidance if they are unsure of what information a particular input-field requires, the app will provide a description. However, ease of use and less time training a team required by a site director are small benefits compared to the possibility that exists if the FAIMS Mobile Platform is widely adopted. Ross, et al (2013) explains the nature of open source software, in that,

“...Open-source principles inform our approach to implementation of mobile data recording software at individual archaeological projects. In particular, our mobile data collection software lends itself to sharing and improving field recording methods and practices themselves – not just the underlying software – using distributed, peer-based production” (p.123).

The potential of the FAIMS Mobile Platform goes beyond simple sharing of research design. It extends to capturing field recording workflows, “...revealing much about the methods and practices used on any given project – an outcome that would contextualize the data and interpretations produced by that project...” (p.125). Detailed recording of a specific project’s methodologies is a crucial component that lends to the context of a dataset. With a digital collection platform like FAIMS, the software inherently captures the methodologies used and extends to the ability of researchers to interpret and reuse the data. Additionally, it provides a more robust picture of a dataset for the purpose of preservation.

The FAIMS Mobile Platform, even in its infancy, is playing a major role in the future direction of the archaeological information ecosystem. According to Shawn Ross, et al (2015) in their article “Creating eResearch tools for Archaeologists: The Federated Archaeological Information Management Systems Project,” “No information management system has yet been created which allows

archaeologists to shepherd their data from digital creation through editing and analysis, to online archiving and dissemination” (p. 114). The FAIMS Project, specifically the FAIMS Mobile Platform, is an innovative way that addresses the gap in the information ecosystem. It provides a starting point in the archaeological workflow where data exists in a digital space upon creation. Additionally, “At the other end of the data lifecycle, we are contributing improvements to a proven online repository-tDAR-which will enhance its capacity for data sharing, allowing archaeologists to derive more utility from their own and others’ data” (p. 114). The capability of an enhanced capacity for data sharing will soon be expanding to include AAI’s repository, Open Context, in addition to a number of other already-established archaeological databases. Sarah Kansa mentions, “We’re working with FAIMS on a data transfer process so data collected using FAIMS can be exported directly to Open Context. We’ve had a lot of people over the years ask us about field data collection tools that can easily export to publication/archiving, so we’re thrilled to be working with FAIMS” (personal communication, October 22, 2015). The questions Sarah Kansa has received in regards to digital field data collection tools indicates that not only is FAIMS gaining traction as a reliable platform to collect data, but that FAIMS is also fulfilling a recognized need in the field in regards to providing open-source, customizable digital tools for initial data collection. As Ross, et al (2015) further elaborate,

“We are not looking to build a ‘walled garden’, but hope to establish a system that includes choices at every stage of development. Many of the components will, we hope, ultimately be developed externally to FAIMS but ‘federated’ so as to allow the automated exchange of data...Use of open, widely accepted and well-documented standards and file types throughout the project ensures that the ‘hooks’ are in place to allow the federation of new components in the future” (p. 114).

The flexibility offered in regards to data collection and use with the FAIMS Mobile Platform, in addition to its capability to allow archaeologists to federate their data to an archaeological repository of their choosing, establishes FAIMS as an integral part of the archaeological information ecosystem. The ability to adhere to all archaeological needs in regards to the type of data collected and how it is collected, then to send that data to a repository archaeologists may have previous experience working with, creates a standard approach handling digital archaeological data.

Conclusion:

The Future of Digital Archaeology

Eric Kansa (2012) discusses many of the financial, academic, and political forces that impact the way archaeology is conducted, even on a global scale, in his article "Openness and Archaeology's Information Ecosystem." Kansa states that, "Archaeology faces the task of understanding and preserving the record of the past with fewer resources and people than enjoyed in previous decades. To meet this challenge, archaeology needs to enhance its productivity and efficiency" (p.504). The task of enhancing productivity and efficiency should begin before the first shovel meets the ground. By utilizing a digital method of data collection, like the FAIMS mobile platform, archaeologists can further streamline their workflow, and eventually enter their digitally-generated field data into the open-source online archaeological information ecosystem.

Although creating one specific standard method of on-site data collection may not be the solution to the current needs and challenges facing the field of archaeology as a whole, it is still imperative that the conversation of standardization remain at the forefront of the field. As Jeremy

Huggett (2012) discusses in his article, "Lost in Information? Ways of Knowing and Modes of Representation in E-Archaeology,"

"Without standards, much of what we take for granted would cease to work. As a result, many of these standards are largely invisible, unconsidered at least until they stop working or have to change...Most of us do not know, and do not need to know, about the operation of these kinds of standards: however, standards which impact within archaeology, whether generated externally or internally, ought to be of interest to archaeologists, all the more so if they have become so customary as to be largely invisible" (p. 541).

The standards that Huggett references here are primarily computer-related and are important to transitioning the conversation to archaeological projects, especially as the field of archaeology moves more toward the digital and further from the analog. Using a set of standard methods of data collection in the field, even if only through the reuse and adjustment of existing digital data collection modules in the FAIMS mobile platform, leads to an awareness of and explicit interaction with standards. The widespread adoption of a comprehensive and customizable digital data collection method can create a "back door" (personal communication, A. Sobotkova, November 5, 2015) into the realm of standardization, where semantic mapping stimulates convergence in practice through the easy editing and reuse of modules in their open source environment. Progressively standardized method of data collection remains invisible at the onset and converges when digitally generated datasets conform to the requirements of existing archaeological databases. This is especially true with repositories that work hard to keep their platforms operating at the highest level of accepted archival standards and practice for the purpose of preserving and reusing the submitted data that they become the stewards of.

Sarah Kansa states, "We also hope to see continuing innovation and experimentation with digital data, as well as theory and discussion. We see this as a new area of research that should get

continuing discussion, not a problem to solve (build an archive and check that task off the list)”

(personal communication, October 22, 2015). It is important to remember that in a continually evolving field, it is still imperative to have discussions regarding the specific needs of archaeologists and creating platforms that can evolve with those needs. When Adela Sobotkova was asked how education can play a role in continuing conversations and the adoption of platforms, like FAIMS, she stated,

“Building IT competence and teaching archaeologists about the problems of technical debt are a way to start. While formal seminars might help, there is nothing more instructive than experiencing the frustration of having to correct incompatible, erroneous data, years after fieldwork. For that purpose and to assist with learning the digital tools, it does not matter what medium the training is delivered through, but it is essential that it is embedded in a social collaborative setting. Learning within a social setting has better outcomes, and collaborations, especially among archaeologists with a bit of interest in IT will have enough energy to carry the group through the learning process successfully. Solitary learning will succeed only with highly motivated individuals” (personal communication, November 5, 2015).

Just by having conversations, listening to and observing the needs of archaeologists, understanding what will enable the best outcome for preservation and reuse of data, and collaboration with others who have similar goals, organizations like the FAIMS Project, AAI, and Digital Antiquity are paving the way for the future of digital archaeology.

Resources

Atici, L., Kansa, S., et al. (2013). Other People's Data: A Demonstration of the Imperative of Publishing Primary Data. *Journal of Archaeological Method and Theory* 1(3), 1-19.

Alexandria Archive Institute. (n.d.). Retrieved from <http://alexandriaarchive.org/>

Digital Antiquity. (2012). Retrieved from <http://www.digitalantiquity.org/>

FAIMS Project. (2014). Retrieved from <https://www.fedarch.org/wordpress/>

Faniel, I., Kansa, E., & et al. (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: ACM.

Huggett, J. (2012). Lost Information? Ways of Knowing and Modes of Representation in E-Archaeology. *World Archaeology*, 44(4), 538-552.

Jacobs, P. & Holland, C. (1997). Archaeology Online: New Life for Old Dead Things. *First Monday*, 2(6). Retrieved from <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/532/453>

Jacobs, P., & Holland, C. (2007). Sharing Archaeological Data: The Distributed Archive Method. *Near Eastern Archaeology*, 70(4), 197-200.

Kansa, E. (2012). Openness and Archaeology's Information Ecosystem. *World Archaeology*, 44(4), 498-520.

Kansa, E. & Kansa, S. (2014). Data Publishing and Archaeology's Information Ecosystem. *Near Eastern Archaeology*, 77(3), 223-227.

Kansa, E., & Kansa, S. (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(1), 88–9.

Kansa, E., Kansa, S., & Arbuckle, B. (2014). Publishing and Pushing: Mixing Models for Communicating Research Data in Archaeology. *International Journal of Digital Curation* 9(1), 57–70.

Kansa, E., Kansa, S., & Watrall, E. (2011). *Archaeology 2.0: New Approaches to Communication and Collaboration*. Location: Cotsen Institute of Archaeology. *Cotsen Digital Archaeology Series*.

Retrieved from <http://escholarship.org/uc/item/1r6137tb>

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

Richter, A. (2014). So What is Digital Archaeology? *Popular Archaeology*. Retrieved from <http://popular-archaeology.com/blog/adventures-in-digital-archaeology/so-what-is-digital-archaeology>

Ross, S., Ballsun-Stanton, B., & et al. (2015). Building the Bazaar: Enhancing Archaeological Field Recording Through an Open Source Approach. *Open Source Archaeology*, 111-129.

Ross, S., Sobotkova, A., Ballsun-Stanton, B., & Crook, P. (2013). Creating Eresearch Tools for Archaeologists: The Federated Archaeological Information Management Systems project. *Australian Archaeology*, (77), 107-119.

The Digital Archaeology Institute. (n.d.). Step 1: What is Digital Archaeology? Retrieved from <http://www.digitalarchaeology.org/what-is-digital-archaeology>