Motivations and Expectations for Central Database Creation at an Egyptian Archaeological Site

By Jasmine Smith

Master's Thesis Committee:
Andrea Thomer, Assistant Professor
Elizabeth Yakel, Dean for Student and Academic Affairs

A master's thesis submitted to the faculty of the University of Michigan School of Information in partial fulfillment of the requirements for the degree of Master of Science in Information 2019

Acknowledgments

Thank you to all of the participants involved in this research for their valuable contributions. Special thanks to Dr. Andrea Thomer, Dr. Elizabeth Yakel, and Dr. Rebecca Frank for their mentorship throughout this research process. Thank you to my family and to God.

This research was partially supported by a research grant offered by the University of Michigan School of Information Office of Academic and Student Affairs.

Table of Contents

Introduction 3
    Purpose of the Study 4
Literature Review 5
    The Field of Archaeology 5
    Intersections of Archaeology and Information Science 5
    Databases and Archaeology 7
    Research on Data Curation and Sharing in Archaeology 8
    Elements of Databases Useful for Archaeology 9
    Implementation Issues 10
    Databases as Tools and Social Impacts of Databases 10
Methodology 11
Findings 14
    Access and Time 15
    Legacy 18
    Integrity and Professionalism 19
    Transparency Among Team 20
    Additional Themes 23
Discussion 28
    Archaeology, Science, and the Role of the Database 28
    Review of Themes 28
    Implications for Information Scientists, Archaeologists, and Information Technology Specialists 29
    Findings of Interest for Further Exploration 31
Conclusion 34
References 35
Appendix 38

Introduction

With increasing advancements in technology, archaeologists are adapting new methods and equipment to supplement or entirely replace traditional, paper-based methods of data collection. For example, some excavations are replacing traditional paper field notebooks with electronic tablets to automatically record and upload site data online. As archaeologists seek new ways to collect, store, and visualize data, they also find themselves dealing with issues of data curation and management during excavations and surveys. To address these issues, some archaeologists, with the help of outside specialists, have created databases to store and manage data for specific archaeological sites.

Though archaeology teams all over the world have implemented databases for their work, the literature on this topic from both an information science and archaeology perspective is limited, especially focusing on Egyptian archaeology. While several archaeologists have published papers about their databases post-implementation, this study seeks to examine archaeologist's needs and motivations before a database is designed. In this paper I explore the motivations behind the upcoming creation of a central database for an archaeological excavation in Egypt, and the impact the team believes a central system will have on their work. Currently, various members on the team use their own systems for curating data, ranging from commercial databases to excel sheets. With an information science perspective, the team hopes to bring their various systems together to form one database that is usable for all team members.

This archaeological excavation produces various types of data including photographs (photographs of stratigraphy, features, and objects), conservation treatment records, artifacts, geographic information, and excavation/site data. Though different specialists on the team deal with specific types of data, they find themselves needing to reference each other's work. This can be difficult as there is no infrastructure in place to see everyone's data, and team members work in different areas when on site, and live in different countries when offsite. Therefore, this study aims to understand the team's information needs and why they believe a central database is necessary.

The research questions for this study ask:

- What are the main motivations for a central database?
- How does the team think a central database might affect data collection and dissemination?

Through these questions I plan to understand the effects of having separate databases and spreadsheets for individual team members on the team's work and why the team thinks a central database will be an improvement. These questions are explored through interviews with the individual members that make up the excavation team. Examples of their datasets and databases were collected as well. Interviewees include the team director, conservator, epigrapher, bio-archaeologist, GIS specialist, ceramicist, small finds object specialist, museum registrar, and photographer. Interviews involved a walkthrough of what the team is doing now and what they want to see change, how a central database will facilitate that change, and what they envision this database should look like. The interviews are semi-structured, with some questions tailored to the interviewee's role on the team. The interviews were also transcribed and coded using NVivo to find themes related to motivations and anticipated impact of a central database. In this study, current data collection and storage processes are examined, as well as the decisions archaeologists make regarding tailoring a database to a site, selecting what data to include, how to describe this data, and how context of data can be preserved.

Purpose of the Study

The research goals of this study include discussing issues at the intersection of information science and Egyptology/Egyptian Archaeology, understanding what motivates the creation of a central archaeological database at an Egyptian archaeological site, and how archaeologists perceive a database to affect data collection and eventually dissemination of information.

Overall this research is intended to:

- Add to the discussion on data management as the archaeology field becomes increasingly digital
- Give suggestions from an information science perspective on data curation at Egyptian archaeological sites
- Be an insight to Egyptian archaeologists on how others in the field are thinking and dealing with the topic of data curation at their site

Undertaking this research is important as Egyptian archaeology produces massive amounts of data that helps inform our understanding of world history. This data is used in a variety of formats including publications that inform the public and academic community of new insights into the ancient Egyptian culture. With the help of information scientists, Egyptologists and archaeologists can find more effective ways to curate data that positively impacts their research goals and the dissemination of archaeological information to the world.

Literature Review

The Field of Archaeology

Archaeologists study people, primarily of the past, through the excavation of places they manipulate. Archaeological excavations are usually conducted by teams consisting of members trained in diverse specializations. These specializations might include a photographer, bioarchaeologist, zooarchaeologist, epigrapher, conservator, registrar, objects specialist, and GIS (geographics information systems) specialist (Faniel et al., 2013). The purpose of these excavations is to collect data that helps facilitate the study of human activity at a site. A site is of archaeological significance if it has a concentration of artifacts or features that signify the location of past human activity (Sagrario, 2013). All of the artifacts, features, faunal and flora remains, architectural structures, etc., are considered data and are recorded and described in a variety of ways including measurements, photographs, drawings, and written notes. This data is seen as a record of human life and is studied to understand how people lived. Members of an archaeology team who specialize in these specific data are responsible for the recording and description of that data type (Faniel et al., 2013).

Depending on the project, archaeological teams have members that come from different institutions and even countries (Faniel et al., 2013). It is not uncommon for some teams to only see particular members in person during the excavation season. Excavation seasons typically run from a few weeks to a couple months. Once the season is over, reports are created based on the data collected in the field. Team members return to their respective institutions and work on publications that inform the archaeological community and the world of the conclusions they made about the human past based on their data. (Federici, T. and Braccini, A., 2012).

Intersections of Archaeology and Information Science

Archaeology projects produce massive amounts of data (it is not uncommon for a single excavation season to yield thousands of artifacts) all of which need to be curated and preserved for a variety of reasons: for future excavators that join a well-established archaeological project, for the team itself to find information and work more effectively with external partners, and to continue to add to our knowledge of the human past. Many archaeological projects return to a site year after year and thus have multiple seasons each with their own mass of data (Sagrario, 2013). Managing this data can cause complications when storing large amounts of archaeological material and

creating descriptive metadata (Gidding & DeFanti, 2014). The material archaeologists uncover, along with details they record about this material, are potential sources of information to better understand the activities that took place in a particular location (Federici & Braccini, 2012). Many in the field are realizing the need for collaboration with information science professionals to find ways to better curate data at their site. On the information science side, scholars admit that archaeological data management and database use within archaeology has limited coverage or theoretical analysis in the field of information research, and that information pertaining to archaeological material is often not managed with proper information systems (Federici, T. & Braccini, A. 2012; Labrador, 2012). The necessary collaboration between these fields is expressed to the archaeology community by Labrador (2012) who states, “It is time that we join our forward-thinking colleagues in information science, library science, museum studies...and to collaborate with those who have taken experimental steps toward implementing new ontologies and data models” (Labrador, 2012). Vincent et al. (2014) express a similar sentiment especially when addressing issues with the longevity of archaeological data, stating that, “The only solution is to tap into permanent infrastructures, ideally libraries that have a strong focus on digital curation” (Vincent et al., 2014).

Managing archaeological material takes into account the various activities performed on the material (e.g. excavation, restoration, conservation, description, exhibition) and though information is needed for each one, it is not always managed in useful ways (Federici, T. and Braccini, A. 2012). The activities that take place on this material are often done in different places, at different times, by different team members (Federici & Braccini, 2012; Faniel et al., 2013). Therefore, team members need to be able to record data separate from other members of the team (Federici & Braccini, 2012). As a result, unless a central system is in place, team members record their data in their own way (e.g. excel spreadsheets and various database software), and these data may not be immediately accessible to other team members. This can be an issue as sharing and managing archaeological data is important for assessing its potential as a key source of information (Federici & Braccini, 2012; Faniel et al., 2013). According to Federici and Braccini (2012), the differing recording systems also make data transferability and interoperability more difficult as different members’ specific systems have different information requirements. Once an object is found it starts a new life that can consist of conservation, storage, or even exhibition in a museum, and each stage of its new life generates documentation (conservation records, accession records, field notes, etc) that need to be managed and archived (Federici & Braccini, 2012).

Databases and Archaeology

A database is a collection of information that is organized and stored on a computer (La Bella and Nagle, 2014). Dictionary.com defines a database as a “comprehensive collection of related data organized for convenient access, generally in a computer” versus a spreadsheet which they define as, “a worksheet that is arranged in the manner of a mathematical matrix and contains a multicolumn analysis of related entries for easy reference on a single sheet”. According to Hine (2006) databases are examples of information technology that “provide a means of gathering together vast amounts of data” and that perform “investigations on those data”. Those participating in scientific work have generally become attracted to databases as places to store and organize their results and act as a shared resource and research tool. Over the past few years large, usually open source repositories have attempted to solve the issues of access and sharing of archaeological data such as tDar and Open Context. This study does not focus on large repositories featuring data from various projects, but instead focuses on databases built and tailored for specific archaeological excavations. Archaeological databases often record several levels of description for a single artifact (Gidding & DeFanti, 2014). For example, a database could include a description of the artifact itself, a description of the position of the artifact when it was found, a description of conservation treatment performed on the artifact, etc. As mentioned previously, these different levels of description and management are often the responsibility of differing specialists on the team, meaning different data is located in different places (Bria & DeTore, 2016). Federici and Braccini (2012) note that the varying locations of data can be an issue when trying to identify material, especially if the person responsible for keeping and managing that material is not immediately available. However, one of the advantages of having a central system is the “increased opportunity to keep and organize information previously dispersed across various locations” (Federici & Braccini, 2012).

One goal for the creation of site-specific databases tends to be forming a system that meets the specific needs of all team members (Federici, T. and Braccini, A. 2012). Another goal is the need for long term storage of physical and digital data, and the need to publish curated data (Gidding & DeFanti, 2014). Many goals of these systems also deal with improving data analysis. For example, Gidding and Defanti (2014) list one purpose of their database is to allow archaeologists to make connections between the material they find in order to develop the scholarly interpretations that contribute to archaeological inference. Databases that allow for data searching and sorting improve archaeologists’ ability to “make more informed decisions and robust interpretations” while in the field (Bria & DeTore, 2016). Archaeologists also seek to implement databases into their projects in order to manage data relationships and allow for these relationships to be easily represented (Gidding & DeFanti, 2014; Bria & DeTore, 2016).

Gidding and DeFanti mention their team's idea of a useful database included the ability to give researchers access to all the data archaeological projects generate to allow for complex analyses.

Another higher-level goal some archaeologists believe databases will help them achieve is providing assistance (specifically to the supervisor or director of the project) with summarizing excavation results. Gidding and DeFanti (2014) specifically note the benefit to the supervisors of the excavation by having a multi-user, integrated, database stating these databases allow for the pulling together of data relevant for their reports and other publications. Vincent et al. (2014) also mention having a central repository allows supervisors to review daily excavation data in one place without having to search through multiple excavators' notebooks. Databases and their usefulness in report writing is also mentioned by Labrador (2012) who notes the ability to search for data points and data aggregates in creating summaries. He states database use is a "necessary waypoint between fieldwork and publication (Labrador, 2012). Vincent et al. (2014) specifically point to their database's ability to give them access to both past and present season data without having to deal with photocopies of field notebooks from previous seasons, which ultimately allows them to produce their final reports quicker. By pulling together data from various members of the team, supervisors and directors can meet their responsibility of reuniting excavated features and material and placing them in a historical narrative that helps the world understand the past activities that brought them into existence (Poehler, 2016).

Some databases are also implemented with the purpose of eventually serving as a research or study tool to assist in finding links between archaeological assemblages (Hein & Kilikoglou, 2011). With the goal of implementing databases to assist in archaeological scholarship, Labrador (2012) studies how databases do not just help facilitate archaeological interpretations, but structure archaeological interpretations and the way archaeologists work. Labrador (2012) states that these databases are artifacts themselves of archaeological knowledge production, and that they are also "archives of former states of being". The former states of being refer to the fact that archaeology is a destructive science - once a site is excavated and disturbed, its original context cannot be recreated (Sagrario, 2013).

Research on Data Curation and Sharing in Archaeology

Data curation is the active management of data throughout the data life cycle (Yakel., 2007; Palmer et al., 2013). Information scientists have begun research on data curation practices in a variety of fields. For example, Fearon et al. (2010) have extensively studied the ways astronomers manage, share, and use data. Their research on

astronomy data gives the information science field an understanding of how astronomers approach data curation tasks. These tasks include making decisions on what data will be used in the future, who will organize and describe datasets, and figuring out incentives and disincentives for others to curate data (Fearon et al., 2010). Research shows that scientists recognize they do not always know how to efficiently curate the massive amounts of data they collect and are looking to data archivists and curators for help (Akmon et al., 2011). Borgman (2012) discussed the urgency of the reuse and sharing of scientific data noting that “science is data and that data are science, and thus provide for, and justify the need for the support of, much-improved data curation”. Research on site-based data curation by Palmer et al. (2017) provides an example of more focused data curation concerned with practices at “scientifically significant sites”. They specifically used “stakeholder analysis and investigation of data artifacts” to study practices at a hot spring geobiology site and developed a framework for the collection and description of this data.

Issues around archaeological data curation and data reuse have recently become the topic of several studies by information scientists. These studies range from data management issues to the needs of archaeological data reusers and recognize the higher level issues that contribute to challenges with managing, sharing, and using data such as technology limitations and cost (Faniel et al., 2018). Faniel et al. (2018) discuss how there are not always incentives or motivations for archaeologists to improve data management at their sites, and that it is not always top priority. They also mention that database design is not always a focus of professional archaeological discussions, and even though they believe preserving their data is important, the skills, time, and money necessary are not always at hand.

Elements of Databases Useful for Archaeology

Some aspects of databases designed for archaeological sites include traceability of events, ability to build a network of relationships with archaeological material, and security (Federici & Braccini, 2012). Web based databases allow multiple users to access and edit data in real-time (Vincent et al., 2014). However, sites in rural areas may experience connectivity issues. For the OpenDig system, the team acquired a temporary local server in order to synchronize data across their devices. The OpenDig framework allowed for data consistency by constraining the data entry system (Vincent et al., 2014). Bria and DeTore (2016) also discuss the consistency their database allowed for which was greatly needed as new team members were joining the excavation.

As for database structure, relational databases that allow data to be linked to data in other tables is more useful than hierarchical structures (Hein & Kilikoglou, 2011). Some scholars also recognize the ability for relational databases to decrease redundancy by linking an infinite number of attributes to a context or artifact record (Bria & DeTore, 2016). Other abilities appreciated in a database are flexibility and adjustability of data formats and fields (Hein & Kilikoglou, 2011; Federici & Braccini, 2012) which speaks to Labrador's notion that "archaeology data are forever incomplete" (Labrador, 2012).

Implementation Issues

One issue with implementation of databases designed by information technology (IT) professionals for archaeologists are system limitations and constraints that cause archaeologists to abandon the system even though they realize there is a need for some type of software for their work (Federici & Braccini, 2012). A takeaway from the process of database creation for an archaeological site discussed by Federici & Braccini (2012) is the process the IT professionals took while developing the database. The lack of interaction with the archaeologists who were the end users of the database affected the aspects of the system the technology experts focused on in the design and development process. Instead of hearing directly from the end users about their needs and wants for the system, the IT experts based their decisions on documents available to them from the team. As a result, the product they delivered was "too bounded for final users" and was seen as a hindrance instead of a tool to help improve the team's work. The tech experts made major changes when redesigning including working interactively with the final users to design the system, and not making standardized forms, but fitting forms to the needs of the different users on the team. This was done by speaking with different users and creating a "wishlist" of system features and a "common thesaurus to be used for labels, texts and voices in the drop-down menus" (Federici & Braccini, 2012).

Databases as Tools and Social Impacts of Databases

From the studies mentioned previously there is a theme researchers found where databases were often seen as a tool for research and for organizing their work. Hine (2006) studied how databases provide new forms of communication and collaboration to scientific researchers. Key issues raised by databases include changes to the ways research is evaluated and to the ways scientists do their work. Hine (2006) notes the differences in how databases bring about change: some past literature mentioned changes in "processes and outcomes of research" where more frequently people discussed changes to the "dissemination of research results and protocols".

There is also discussion on how using databases for science suggests the idea of knowledge being “altered by new representational technologies”. Hine (2006) also notes that the databases are not independent agents of change and that they need to be “embedded in an appropriate set of work practices”. With that being said, this study asks a specific team of scientists how they expect a central database to cause changes in their work.

Methodology

This paper is framed as a case study and seeks to understand one archaeology team’s motivations and expectations for future implementation of a central database for their project. This team was identified for this study through attendance of the 2017 American Research Center in Egypt Annual Meeting, where I spoke with Egyptologists about what information science related-issues exist in Egyptology and Egyptian archaeology. At this meeting, the director of the project featured in this case study mentioned their team’s desire for a central database. The team conducts archaeological field work at a tomb site in southern Egypt, west of the Nile River. The site is part of a larger area home to burials of important officials and royalty from the early periods of ancient Egyptian history where multiple teams hosted by various universities and other institutions currently conduct excavations. The team mentioned in this study are focused on a tomb belonging to a government official who lived during the Old Kingdom period of ancient Egypt (ca. 2750-2260 B.C.). The site consists of two primary burial complexes and stretches about 80 meters wide, east to west.

Excavations of this part of the site have occurred prior to the start of this team’s project, some starting in the mid-19th century. The excavations that took place in the 19th century resulted in the discovery of inscriptions that have been important in understanding ancient Egyptian political history. However, the excavations were run by explorers primarily interested in finding artifacts to fill their museums, rather than producing detailed publications. As a result, descriptions about the burial the inscription was associated with and its exact location were unknown. The location of the site was rediscovered by the current team in the mid-1990s as they began studying objects in museums that originally came from the site, and performing archaeological surveys. The team now returns to hold either excavation or study seasons each year. Excavation seasons involve actual digging activity, while study seasons involve no digging, but study of data found in previous seasons.

There are a variety of sub-teams for the project that focus on specific areas of work. These include the excavation team consisting of those actively digging and recording archaeological data, the bioarchaeologists who excavate and record human remains data, epigraphers who study and record inscriptions on stone, the registrars who are responsible for tracking all artifacts excavated each season, the photographer who captures images for any aspect of the project deemed necessary by the director, ceramicists who focus on the study of all ceramic artifacts such as pottery, the conservators who treat and evaluate the condition of objects, and geographic information systems (GIS) specialists who capture the excavation's spatial data. The project is based at a large midwestern university, however, most team members are based in various continents including North America, Australia, and Europe.

In developing this case study, I use qualitative methods to understand archaeologists' motivations and expectations for creating a central database. These methods include semi-structured interviews held with 9 individual members of the archaeology team over a four-week period. I chose to use qualitative methods in the form of interviews in order to gain an in depth understanding of each participant's role on the team, the data they work with, their research goals, and their information needs. These interviews were primarily conducted virtually using Zoom or Skype and were audio recorded. Interviews ranged from fifteen minutes to an hour and a half.

An interview protocol was drafted and used to structure each interview. Interviews started with questions about the participant's role and professional history at the field site. These questions were asked in order to get a sense of who the participant was and what data they collect and/or manage. A second section of questions was dedicated to data collection and management. In this section I asked participants to walk me through their process of working with an object, excavation unit, etc., and what information they record about it. Participants also described the databases or spreadsheets they use if any, why they started using them, and what they like and do not like about them. The final sections asked participants to explain the factors that went into preliminary discussions to create a centralized database system, explain how they envision the system being used, and to describe what technology and information specialists need to know about how the team works in order to build a useful system. The interview protocol is available in Appendix 1.

The audio recording of each interview was transcribed by a transcription service. The interview transcripts were coded using NVivo and the codebook was developed using open and axial coding (DeCuir-Gunby et al, 2011). During the open coding process, I developed the codes based on the research questions and themes that appeared during interviews. Open coding is the process of "breaking data apart" and assigning

codes that represent a block of data (DeCuir-Gunby et al, 2011). Open coding was used in this study to understand what (if any) themes appeared throughout the interviews. The axial coding process involved analyzing the codes in order to identify any connections between them. The connections between codes founded the major themes which will be discussed in the findings section.

Findings

As mentioned previously, the team's current data practices involve sub-teams managing data separately from one another. Different sub-teams and specialists use their own databases, spreadsheets, and notebooks to keep track of their work. Several team members spoke of databases the team used in the past and the relationships between these systems and their work. Team members also discussed the differences in technology and using these systems in the early years of the project versus the present. One team member stated,

> *...initially there was a FileMaker database, but it was hard for everybody to use it. Especially in the early days because we couldn't walk in there, everybody clutching a laptop like we do now. There would be one or two designated computers, because things like that during those days had to be registered coming in and out of the country. (mtop2)*

When databases proved hard to use or were unavailable, some team members took it upon themselves to find something that worked for them. For example, one member explained, *"I put together a database just for the stone, because otherwise it would just be sort of a jumble, nightmarish mess"* (mtop1). Another member spoke about archaeologists in general and their attempts to build databases explaining that the use of Filemaker by archaeologists was not a result of it being an "amazing database program", but because it could be used and manipulated without having to know much about databases (mtop8). However, issues arise when multiple people on the same team create databases that are isolated from other areas of work on the project. These databases are not relational and are structured much like spreadsheets, without connecting to the work of others. One member stated, "*A lot of archaeologist's databases are like that. When you go in and try to actually look at how the fields relate to each other...it's just not [relational]*" (mtop8).

In coding, four main themes regarding my participants' motivations and expectations of a central database emerged: access & time, legacy, transparency & context, and integrity & professionalism. Therefore, the rest of this section is divided into these themes with subsections describing the motivations and expectations related to each. Other findings of interest not directly related to these themes are discussed in the additional themes section. In general, team members expressed that the point of creating a central database is 1) to make their information usable, 2) so that anyone looking back at their excavation information would know that this site and project existed, and 3) to organize a 20 year backlog of data and ensure information recorded in the earlier, pre-technology days of the project is preserved.

Access and Time

Motivations in creating a database related to Access and Time: Several members on the team said they are motivated to implement a central database because accessing each other's data (and sometimes their own) takes an inconvenient amount of time. This is a major problem because time is one of their greatest constraints, especially when at the field site. The current methods of organizing these data and making them accessible to other team members has proven to be a challenge. One member described the situation as,

> *...a very messy closet where everything's shoved in. We don't even see the shelves and the hangers anymore. It's just all shoved in the closet, and all of a sudden we want something and we don't even know where it is. (mtop2)*

This member also described the process and how it is problematic, saying, *"...how the heck do we get all that (information) without spending three days going down the rabbit hole of sentimental file searching?*" (mtop2). This team member is explaining that while looking for specific information, one often becomes distracted by files unrelated to what they are looking for. These distractions caused by lack of direct access compound the problem as team members spend more time trying to access information.

Not only is locating one's own information a problem, but so is retrieving information from other team members. The use of many different information organization systems (systems being defined as whatever they use to store data, *e.g*. a database, spreadsheet, etc.) increases miscommunication and is not ideal for research or later data analysis. Given the international nature of the project, and the way in which team members come and go from season to season, some feel it is important to have a central system everyone can access, instead of holding on to their data and keeping it to themselves. One team member emphasized this saying,*"…the research doesn't belong to me, it belongs to the project and it belongs to Egypt"* (mtop1). The team is thus motivated to move toward a solution that allows them to share data with each other and cut down the time it takes to do so.

Team members expressed confusion about finding information for excavated objects, and some believed it is better to have all information related to an object in one place. Having multiple databases that are not interoperable has caused some members to miss object information due to lack of communication between systems. Several team members stated that keeping information about objects in one place that can be retrieved at one time would not only cause less confusion, but be more useful for future research purposes, such as object interpretation and analysis,

For any kind of future analysis or interpretation, you really need all of the material in...one consistent framework...it's kind of nuts to have all kinds of different systems, because then things fall through the cracks and it's not accessible to other people. (mtop4)

Out of the several constraints the team deals with (which are discussed later in the Fieldwork Constraints section) time is often the most pressing. The time the team has to work at the field site is limited – just a few weeks each season – therefore team members are consumed with getting as much done in the field as possible. Because they do not presently store data in one accessible database, team members spend more time than they would like on trying to access information.

Several team members described problems that occur when having to access information stored in other members' systems, particularly when that member is unavailable. These problems included trying to access a team member's data who did not return to the site during a season; contacting international team members who might be in completely different time zones and having to wait for a response; or having to track down members while they are working on different areas of the site. Problems such as these hinder each other's progress in their work, for example, writing strong funding proposals:

Not everybody turns their reports in, in a timely way. For people whose reports are really contained in the database that they're using, if "X" doesn't have access to that database, then "X" doesn't really know what they're finding and what they're doing. There's no malicious intent behind that, it's just people being busy and not turning stuff in. But it means that sometimes you don't have access to information when you need it. You need to write a grant proposal and you don't really know what's been happening with some aspects of the project. (mtop8)

Some team members noted that publishing about the project can be challenging, partly because of the time it takes to sort through data. One team member said they believed better organization of data would improve this:

When you're in the thick of excavating and being a professor...It's hard enough to keep yourself analyzing and publishing as it is, but having your data tidily sorted, I think, will make it easier to make better headway on it. (mtop9)

Several team members noted the relationship between time, access, the practice of passing around information while on site, and the amount of time this takes:

When we're in the field, it [data] lives on our hard drives and in USB's that we basically are passing back and forth... In a perfect world, we all will be able to access it... Right now I do spend a lot of time passing stuff to people on USBs like photographs or notes or things like that. (mtop8)

Others said that not having information in a central place also takes a significant amount of time when not in the field because team members are based at many different institutions:

We're a much more geographically dispersed group now. So, rather than having to send somebody a FileMaker version of just the excavation database, and maybe they would have to then convert that to 'Access' or whatever; a central system where everybody is working on the same platform and can look at each other's contributions, regardless of whether they're in Austria or Egypt or up here in northern (state). (mtop7)

Based on these points raised by team members it is clear that there are several motivations for creating a database related to time constraints and access issues. These include wanting to decrease the time it takes to access each other's data and their own data as time is limited and lack of access hinders team members from completing their work.

Expectations for a central database related to Time and Access: Team members generally expect that implementing a central database will make it faster to access data, as everyone's data will be together and readily available. Some members stated specifically that trying to access information from their own files or files of others made the team realize they need a database as opposed to spreadsheets and lists. The types of queries and analysis the team would like to do cannot be accomplished using the current range of spreadsheets and databases used by different team members. Therefore, a central database is expected to alleviate the struggles of locating, accessing, and retrieving the information they need.

For specific sub-teams that create reports of their work, having a database is expected to decrease the amount of time associated with creating reports, as information about objects and the work done to the objects would be in the database, and instructions on how to access object information in the database would be referenced in the report. Particularly for the conservation sub-team, conservators would rather record information in a central database than record them extensively in conservation reports: *" In a perfect world...instead of having all of these things, in this Word document, they would actually be in the project's database"* (mtop8). This team member explained that the other

project she works for has a central database that allows her conservation reports to be shorter, 2-page summaries. Because she can include links to the central database, her reports can simply summarize her work and readers can refer to the database for more detailed information. Without a central database, her reports for this project become 30 pages. According to another team member, the team is behind on publication and expects a central database to be a big step in facilitating this: *"I think the big thing is going to be ease of access to information, and that is going to significantly speed up time to publication…"* (mtop7).

With a central database, team members expect to access the information they are looking for without first contacting other members who will have to find the information and then send it,

> *He's no longer going to have to email me to find a pdf of an old form or a notebook; he will simply have access to what he's looking for, and not have these three- to four-day delays on trying to find information either on a hard drive I've got here, or then having to relay that request to [director's name] to pull things off of a hard drive in [city]. (mtop7)*

Team members generally expect the database to make the project more efficient and save them time by decreasing the need to, *"track people down and get them to look things up"* (mtop4).

Legacy

Motivations for creating a central database related to Legacy: Members of the team noted that time not only works against them when trying to access and find information, but also biologically:

> *Some of it is just the simple biological fact of the matter that the senior staff of this project are all aging...there's no way we're going to finish the publication of everything that we found. Some of it is feeling time passing, knowing that people's health changes, their interests change, their availability changes. (mtop9)*

With this in mind several team members noted that a central database is necessary for work at the site to continue, *"10, 20, 100 years from now*" (mtop5) and are thus motivated to implement such a system for the future generation of team members. The director in particular feels a responsibility to put such a system into place, "*We didn't have the data all centralized anywhere. I wanted to be more responsible about that, and also beat the data into a shape that someone else can continue with it when we're not working on it anymore"* (mtop9). It is clear that several members on the team, especially the director, have thought critically about the future of the project and the project's data.

Understanding that people who collect and keep important data from the site are not always available and will at some point no longer work at the site has motivated team members to ensure everyone's data is in one place.

Expectations for a central database related to Legacy: Several team members specifically noted they expect a central database to be a legacy item of the director's academic career and representation of their contributions to the fields of Egyptology and archaeology: *"[The database] it's a legacy item. It's an amazing site. It's an amazing amount of information that she gets from each season...and when you do archeology it's a one time deal"* (mtop2). A central database is not only expected to serve as a means to pass on data to the next generation of team members, but is expected to preserve the legacy of those who currently work at the site. It is also clear that the team is not only interested in a central database's technical capabilities, but its ability as a representation of the team's contributions to Egyptian archaeology.

Integrity and Professionalism of the Project

Motivations for creating a central database related to Integrity and Professionalism: Among several team members, databases are seen as symbols of professionalism and as a representation of their values (e.g. honesty, trustworthiness, openness). One team member claimed that having one database everyone can use that has a certain layout and understanding is *"very much professional"* (mtop3), while another also stressed that the team needed a central database for the integrity of the project: *"...there's no question about that...the question is that we need a good one. We don't need a spreadsheet. Proper, academic, database. Something that is a tool. Not just a record"* (mtop1).

It is important to the team that the Egyptian authorities not only know what they, a foreign team, are doing at the site, but that they can ensure some level of trust, so that the authorities also know the team is managing the country's cultural heritage data in a responsible and academic manner,

> *I should think that it [the database] should be accessible to the Egyptians [antiquities service officials] to relieve any worries they might have about the integrity of our storage systems and our intentions regarding their objects. So I think it's important that they know that we're working within strict academic bounds, that we are totally trustworthy...the more integrity we can show that we have and the more honesty, the more openness with them, with individuals within the antiquities service, then, I think the better it is. (mtop1)*

The Ministry of Antiquities is the section of the Egyptian government that monitors all archaeological field work in the country. Foreign archaeology teams usually have a representative from the ministry accompany them while on site. From this team member's statement in particular, the team's awareness of their position as foreigners handling extremely valuable and meaningful data is evident. They are interested in any means of ensuring they are caring for this data appropriately and proving that they are committed to the academic study of the country's cultural heritage, especially considering the country's history of looting by foreigners. Thus, maintaining relationships and credibility with the Egyptian people has motivated the team to continuously improve their work practices.

Expectations for a central database related to Integrity and Professionalism: Some team members mentioned specific examples of how a central database is expected to assist in fulfilling requests from the antiquities service, such as the ability to create object lists. One member also discussed how a database will allow for sustainability of the project's data, even if the team were unable to continue work, saying,

> *I think we're responsible and we do a good job but there might be...a level of...confidence or relaxation that would come with knowing that the [project name] project has [a] centralized database and even if something bad happens and [we] can never come back, we [the Ministry of Antiquities] could have that information. (mtop8)*

Here we see the expectation of the database being a source of relief in knowing that the project's data is kept in a central place in case the team were unable to continue work at the site. Overall, from the team's motivations we see that a central database is expected to be a representation of the team's intentions, values, and professionalism.

Transparency Among Team

Motivations for creating a central database related to Transparency and Context: Several team members mentioned issues with the current team dynamic and the isolated nature of their work as motivations for a central database:

> *I think we've been working in our individual areas to a certain degree...of isolation and I don't think that's been intentional. I think it's because we're all so busy... everyone's head down, fast as they can trying to get through what they can do in the time that we have on site because it's always limited. (mtop1)*

This isolation makes understanding what other team members are working on more difficult, and prevents team members from understanding how their individual work fits into the larger picture of the project. One team member said,

> *Transparency is important, so that we all understand what we all are doing. From my point of view, it's frustrating for me to look at the work that I do and see the isolation – to my mind it's part of a whole. This [artifact] is not just a stone pillar, this is a stone pillar that came from a chapel or came from a* mastaba *or came from somewhere on the site that is linked to another building that is linked to other objects. (mtop1)*

Having transparency about what different sub-teams are working on and discovering is necessary for others to understand the context of the materials they work with. This transparency entails better communication between sub-teams so everyone has an idea of each other's goals and motivations,

> *We don't all necessarily have a great understanding of each other's goals and motivations. So, that can sometimes cause discord just because one part of the team wants to do things a certain way and that doesn't really work very well for the person two steps later who needs to actively use the information generated three steps ago. I don't know that having a centralized database would actually help fix that problem, but it might...So maybe it would create a more kind of exciting collegial community. (mtop8)*

Those who have been on the team for several years become accustomed to their own workflows, however, some team members are noticing these workflows might not be the most efficient for the team as a whole. For example, one team member stated, *"they've kind of become familiar with the idiosyncrasies, and we don't realize maybe that's not the best way for the group effort*" (mtop2). Though several team members want a central database that provides a level of standardization in the way data is entered across the project, the inability to let go of current work practices and idiosyncrasies might cause tension between these two desires. More on the team's desire for standardization is included under the "Additional Themes" section.

Individual members of the team understand that their work is linked to that of others, but not having the ability to visualize these links perpetuates the feeling of isolation. For instance, one member said she wants to summon what was found in specific units and excavation seasons (excavation contexts) or even summon objects related to a certain period in ancient Egyptian history (historical contexts), and for data to be searchable by different fields. These query capabilities would assist her in more easily drafting reports about a fieldwork season. One team member described a situation in which the conservators were struggling to treat an object "*unanchored from its contextual information*" (mtop9), meaning they were unaware of the object's historical context and how to approach its conservation. Conservators rely on information from object labels created by the excavation team that give details about where and when an object was

found. Knowing where an object was found helps conservators understand the environment responsible for the object's preservation or deterioration, which allows them to determine what treatment to perform. Furthermore, knowing the historical context of an object is also useful in determining conservation treatment, as conservators can compare how certain materials were used between different time periods in ancient Egyptian history. However, information about an object's historical context is not always immediately known and thus rarely written on object labels. Therefore, the team's conservators are motivated to implement a central database that allows team members who determine historical context to enter this information so it is accessible to those who rely on it for their work.

Expectations for a central database related to Transparency and Context: By storing everyone's data in one system, the team expects to use the central database in a way that helps them see the work carried out by other sub-teams of the project, and understand how their individual contributions are related to those of other team members'. Several team members define transparency as the ability to see what other sub-teams contribute, and the ability to access each other's contributions provides context for other areas of work at the site. As shown in the section above, one team member discussed how the conservators needed context from other areas of the project for their work. This team member then discussed how she expects a central database to facilitate this by stating, *"...if she printed out reports from our central database, 'Okay, give me all the wood' then those reports would ideally show this location, this date, assigned to it, this complex that was associated with other things that came out of the ground with it"* (mtop9). Another team member also discussed the needs of specific sub-teams and described how a database will support their work, saying that it was clear that,

> *...the other non-excavation teams, and individuals within the overall project needed conceptual information from the excavations. So, having a larger database, a data management system for all of the little sub-groups within the overall project, is going to make it simpler for them to get the conceptual information. (mtop7)*

Beyond the expectations there is a desire to understand exactly how a database will allow the team to preserve context. One team member said, *"...context is an invaluable part of this. I guess I don't know enough about database development, how to make that a tighter relationship between context and object"* (mtop9).

It is clear that a central database is not merely understood as a means to store and retrieve data in one place, but also as a tool where the team can make sense of the data in the scheme of the larger project. Team members are particularly interested in

the ability to generate lists of objects that came from specific excavation contexts, for instance, *"...being able to click on a polygon that represents an excavated level in the GIS, and say, 'Hey, let's generate a list of all the objects that were excavated from here'"* (mtop7). These lists are useful for individual team members and also for the director, who is responsible for understanding the overall picture of what is happening at the site. One team member explained this saying,

> *I would imagine that if you are [The director] and you are thinking about a particular place on a site and how it's being used, it would be nice to see not only the excavator's notes and photographs about that place, but also...build out in the database a list of all of the objects that came out of it (mtop8).*

Recalling a time they were asked by a museum to put together a proposal for reconstructing one of the structures the team is excavating, another team member said that, *"pulling together the list of object numbers for that was a nightmare. Instead of being able to go to a database and say: 'X' blocks"* (mtop9).

In general, team members expect a central database to provide context from various areas of work on the site in order to inform their individual tasks such as building reports and performing conservation treatments. Team members also expect a central database to allow for more transparency in terms of what sub-teams are working on and finding, and what their goals are. Finally, a central database is expected to allow team members to understand how their individual work fits into the larger project by viewing their contributions in the context of other team members'.

Additional Themes

Other themes that emerged from interviews include the constraints the team are under, the use of a database as a research tool and tracking system, standardization within databases, and possible issues with implementing a new database into the team's workflow.

Fieldwork Constraints and Challenges

When in the field, the team often deals with certain constraints and challenges including limited time for excavation and artifact handling, language differences, and unreliable internet access. Some of these constraints, such as time were mentioned previously in the "Access and Time" section. For instance, one team member discussed how time restrictions for accessing excavated material means objects occasionally need to be studied digitally. They stated,

Because of the time constraints we have on the site, and also...this year the constraints that will be put upon our working practices [where we] have a certain number of hours where we can access the material in the magazine...so a lot of this can't be done in real time with real stone, it has to be done with virtual and digital... (mtop1).

Some team members also mentioned the physical aspects of being in the field and how these can have an impact on their work, saying,

...if you could see the working situation you'd perhaps understand it's stressful, not just in a lot of people together in one place, trying their best to work together effectively, but also the climate, just the whole environment can be challenging...also working with an international crew. (mtop1)

Working with an international crew can specifically cause challenges as some team members have varying native languages,

Although English is a common language, Arabic is other people's language, German is other people's language, French is other people's language so we're all trying to make sure that everyone understands what everyone else is saying which is why the record keeping...has to be so precise. (mtop1)

These constraints and challenges have motivated the team to put their data in one system so that accessing each other's data becomes one less problem. Given these constraints many team members believe a central database will overall make their jobs easier:

"Yeah, I think it will make it easier for everybody to have access to it [data], is all." (mtop4)

"First of all, it would make our job easier." (mtop3)

"It would be easier for me to find things." (mtop8)

One team member said that a central database would make her job easier because it would give her the ability to create artifact records instead of copying information into Excel spreadsheets. This would allow others to query the database and discover the information they are searching for much easier than "*going through thousands of cells in the Excel sheet"* (mtop3). Another member said that using a database for data entry would be faster than using a Word document and spending time on formatting. She discussed how a database would also make it easier for her to find information regarding the conservation of objects, saying, *"If I remember that I did a treatment or I*

remember that I worked on an object, but I don't necessarily remember what season I did it...theoretically I can search the database and find it" (mtop8). One member describes specifically how the current system of recording data is not ideal and that a database could change this, because,

> *...we just put [data] into notebooks and Excel tables and...I don't want to do that. It [the database] should basically reflect the steps and what we actually do in the field...It [the database] shouldn't only just replicate the steps we do in the field to record, but it should in fact probably make it easier. (mtop6)*

From these statements we see that the team expects a central database to generally make aspects of their work easier given the constraints of fieldwork by replicating how data is recorded in the field and enable them to spend less time formatting documents.

Database as a Tracking System and Research Tool

Several team members said that one of the main motivations for having a central database is to track the vast amount of data collected during excavation seasons. One team member specifically explained that, *"...there [are] so many objects from so many seasons. It's really hard to keep track of them all"* (mtop4). The importance of tracking data and a database's ability to facilitate this is crucial as pointed out by one member who expressed not having a database for their work is equivalent to *"jumping off a cliff"* (mtop1). In other words, approaching their work without an efficient way to track their data is extremely challenging. Not only does the team need to track objects they collect from the site, but also objects collected by others who worked at the site centuries before. Several objects that were excavated from the site before this team's project are in museums around Egypt and the world. Tracking these items is important for the team's research and makes them aware of what has come from the site before their project. Tracking items is also important as it allows team members to compare all known excavated material (whether excavated by their team or not) in order to study the site's history. The team assigns their own tracking numbers to all objects, and some objects are taken by the Egyptian Ministry of Antiquities and given another tracking number assigned by the Ministry. The Ministry's tracking numbers, along with the project's numbers are kept by the project in order to know what material came from the site. Some team members need to track data in specific way; when speaking of previous databases made by team members for use at the site, one member noted, *"It's never thorough enough, because then you forget there's all these other people with their wild things they want to include. Or have gone rogue and made their own databases, because they've got to track it"* (mtop2).

The team expected that with a database they would be able to properly track material more efficiently. One member specifically mentioned the importance of tracking objects in a database as objects are passed from person to person on the team. Objects are passed from the person who excavates them unto others such as the photographer for pictures, epigrapher for inscription study, registrar for tracking, and conservator to perform treatment if necessary. Each stage of passing along these objects involves team members collecting metadata such as size, weight, treatment performed, and description, all of which need to be tracked from person to person.

Similar to the importance of tracking, some team members expect a main function of the database should be the flexibility to serve as a research tool for the team and for outside researchers. One member specifically argued that the point of scientific expeditions is to provide a system to allow people to retrieve information, and that their systems for storing data need to be useful to others beyond the person or people running the project. This person said, *"...if a researcher comes in...and says gosh I'm really curious about X, Y and Z, they can...pull it [data] out and go do their own project"* (mtop2). Some team members mentioned for their own analysis of data they need a system that allows them to easily query and extract the data they're looking for. Another member said, *"It's a question of not being able to easily lay your hands on all the relevant data to a particular question that you're considering"* (mtop9). In general, team members would like to use a central database as a research tool during their specific work on the site such as comparing newly excavated artifacts to those from previous seasons. The database is thus expected to allow the team to evaluate data in order to facilitate study of the site.

Standardization

With a central database team members are expecting to improve the standardization of how data is represented and recorded. One member described specific features for standardization saying,

> *This is why I like drop down menus and things like that because like, oh yeah, I forgot to say 'yes or no'. 'Was it in a coffin? yes or no'...We're all slightly compromised when we're out there [in the field]. Literally. We're tired, we're thirsty, we're under the gun on time and getting things done, so if things are kind of spelled out for us a little bit, we can make sure that information gets checked as absent, present, I don't know, destroyed, whatever. So it [data collection] will change. And it [data collection] will change for the better, because the information recorded will be more standardized. (mtop2)*

By everyone entering in their own data, some believe the data will be more correct, instead of other team members entering information recorded by someone else. Standardizing the language used in the database was also mentioned specifically in recording dates (as team members from different countries record dates differently), spelling the names of other team members, referring to the governing organization (Ministry of Antiquities vs. Supreme Council of Antiquities), and describing excavation elements (unit vs. operation). One member specifically noted the new database should use *"...drop down lists to narrow down mistakes"* and that fields should, "*avoid free text"* (mtop3).

Implementation Issues

Several possible issues regarding the implementation of a central database into the team's workflow were mentioned by team members. These issues included deciding who would be responsible for the maintenance of the database and determining how to link everyone's work. For example, one team member said, "*We have had a few heated discussions at the dig house amongst all of us about what that key field should look like, and I think that's actually going to be the trickiest part"* (mtop7). The key field was mentioned as being particularly tricky due to the idiosyncrasies of the team's nomenclature in describing excavation contexts. Others mentioned the biggest implementation issues will be inputting data from their 20-year backlog and convincing people to use a new system instead of the ones they have used for years. Similarly, some team members noted that in order for the database to be effectively implemented it has to compliment how sub-teams within the team work, *"I think every little sub-discipline within the team is going to have their own quirky wants and their own quirky systems"* (mtop7). One member specifically warns that if the database is not able to adjust to how the team works, its adoption will not last,

> *I think that the way people are entering data really has to work for them...if you try to make it fit into something that they wouldn't naturally use as an end-product or it would be difficult for them, I think they just won't do it. We've worked without it for a long time. So they'll just think the hell with this, I'm just going to keep doing my own thing. (mtop8)*

Discussion

Archaeology, Science, and the Role of the Database

As noted previously, there are aspects of the field of archaeology that need to be understood in order to implement effective tools that meet archaeologists' data practice needs. These aspects include field constraints, work practices, project goals, and how archaeologists view information and communication technologies – in this case, database systems. It is clear that most members of this team view a database as having a larger role than just a means to track and store data. As one member described, a central database should be a *"living and breathing"* part of the excavation (mtop1). A central database is expected to be just as significant a member of the team with its own contributions to the goals of the project as a scientific investigation. According to one member, a central database should make it possible for,

> *Somebody who's never been out there but who has the basic understanding of the site, can get in [the database] and learn about it. Somebody can ask questions and get answers. No more institutional knowledge. It can be used by somebody who hasn't been there. To answer questions. And that's the point of science. (mtop1)*

A central database is expected to serve as a record of the project's work that allows team members and researchers to pull data for reconstructing ancient places and examine how these places have changed overtime. One team member stated, "*…every step of the way [of archaeology] is part of the provenance information of not only things, but interpretations. And if you lose that, then you're just making up a story. You're wasting time and money*" (mtop2). Archaeology is an ongoing process and information about data tends to change over time. For example, *"...a huge site like this, excavation projects are multi-year and something that you label one thing when you first encounter it can end up with a different number by the time you really understand what it is"* (mtop9). Therefore, it is important that database systems are designed to take into account these aspects of archaeological work in order to be effective solutions for this field's data needs. With these ideas in mind, it is necessary to discuss the implications of this research and think critically about the themes behind the motivations and expectations for a central database.

Review of Themes

The four main themes that emerged from interviews with members of this archaeology team show the complexity of how databases are viewed and the roles they are expected to play at a specific field site. These themes include:

- *Access and time*: the lack of easy access to other's data, the time it takes to access this data, and the general lack of time the team has in the field are all issues that have motivated the team to bring their data together. By bringing this data together, the team expects the process of accessing each other's data to be more efficient by replacing the "middleman" between the data user and the data with a central database. It is also expected that this system will serve as a means to find more detailed information about the work of sub-teams which will allow those teams to spend less time on lengthy reports.
- *Legacy:* the team saw a central database as an embodiment of their academic legacies, and expressed concern for the future of the project and consideration for future team members as motivations for a central database. With everyone's data in one system, the team expects to have access to the data of those who will eventually retire from the project, ensure that the data can be passed on to new team members, and for the database to serve as a representation of the director's academic legacy.
- *Transparency and context:* several team members had feelings of isolation while working on site, and the need for contextual information from other sub-teams to do their work were motivations for implementing a central database. The team expects a central database to allow members to more easily share and view each other's data in order to have a better understanding of each other's work. It is also expected to give team members a better sense of how their work is a part of the overall study of this ancient site.
- *Integrity and professionalism*: the team described several values related to their commitment to responsibly handling Egypt's cultural heritage, ideas of databases as symbols of professionalism, and the desire to maintain relationships with the Egyptian people as motivations for a central database. By implementing a central database into the team's workflow, team members expect it to serve as a representation of their values and commitment to continuously improving the ways in which they study and work with Egypt's cultural heritage.

Implications for Information Scientists, Archaeologists, and Information Technology Specialists

Information Science

From the team's motivations and expectations for implementing a central database illustrated by a consensus of four main themes, this case study has implications for those involved in current and future discussions on creation of central, site-based databases. Information scientists in particular benefit from this research by having an example of what data practice issues exist in archaeological excavations and where we can contribute. While studies related to archaeological databases are not new, this case

study provides a clear example of what archaeologists believe databases are and how they have been used (successfully and not) specifically within the context of Egyptian archaeology. Prior work at the intersection of data curation and archaeology shows that these collaborations are necessary, especially when moving from a stage of curation to one of publication and dissemination. Kansa et al. (2014) discussed this as they aimed to build models for communicating archaeological research data saying the process, "required significant investment of effort and expertise, including archaeological domain knowledge and familiarity with key ontologies." According to Faniel et al. (2013), the widespread adoption of, "data documentation guidelines, standards, and ontologies" in archaeology is yet to be seen, perhaps indicating an area where information scientists can conduct further research. Faniel et al. (2013) also found that repository practices in archaeology are more focused on data collection than data reuse. They studied the needs of archaeology data reusers and applied their findings to "existing work on standards development." This work presents an opportunity for information scientists to contribute an understanding of data curation and data standards to the domain of archaeology. As I discussed in the "Intersections of Archaeology and Information Science" section, archaeological data management research is limited in the information science field, and more collaboration between the two fields is necessary in order to develop better data models and ontologies (Labrador, 2012). It is important that more research is done on the specific ideas expressed from this case study such as the views of a central database being a representation of values and legacies. Further research on how archaeologists view database systems and what role they serve helps inform our approach to what effective data curation in archaeology should look like. A further discussion on future research regarding archaeology and the sociotechnical aspects of databases is under the "Findings of Interest for Further Exploration" section.

Archaeology

Archaeologists benefit from this research as it allows those with expertise in data curation and data management systems to understand the needs of archaeologists and the nature of archaeology that needs to be captured in database systems. The ability to capture the nature of archaeology is particularly evident in the team's need to track artifacts located in on-site storage, artifacts taken by the antiquities ministry, and artifacts that were excavated before their time that are now located in various countries. We also see the importance of capturing how specific specialists work and the differences between, for example, how an excavator may enter and retrieve data from a system versus a registrar or conservator. This study also provides an example of how a specific archaeology team is thinking about databases and allows archaeologists to compare what issues and thoughts around databases are similar and different to their own. Through studies on the motivations and expectations of other teams in a pre-

central database implementation stage, archaeologists gain an opportunity to reflect on what databases can and cannot do for their data, and their overall team dynamic.

Information Technology

From this study we see that members do not simply want a central database, but a tool that can be used to understand their work from the perspective of the whole project in order to gain contextual information. Having a better understanding of not just what the team thinks they want in a database, but how they intend to use it, and how they expect this system to impact their work is necessary for those tasked with building such a system. As explained by Federici & Braccini (2012), the IT experts building the database initially bypassed this step which caused the resulting system to be difficult for users to apply to their work. Through studies focused on what motivates archaeologists and what they expect from databases, IT specialists and database administrators have examples of what archaeologists are actually looking for when they say, “I need a central database.” Understanding the range of archaeologists’ views and ideas of what databases are will then allow for better design of database systems.

Findings of Interest for Further Exploration

It is clear from this case study that archaeologists do not view databases merely as places to store the vast amount of data they collect, but as tools for research, communication, transparency, and ultimately as a legacy to their work. This idea of a database as an academic legacy item and a representation of a person's career is particularly interesting and worth being studied in the future across the field of archaeology. These studies help to understand the relationship between archaeologists and databases and how/why they are used. As mentioned earlier it is clear that archaeologists are making site-specific databases, and the literature about how, why, and what decisions archaeologists are making when creating these databases is limited. Further exploration is needed on the systems they are using and why they use them (e.g. as mentioned earlier some archaeologists use FilemakerPro because it is easy to manipulate), what archaeologists expect from these databases (both technically, academically, and socially), and do these databases meet their expectations after they are created? The idea of a central database’s ability to express team values is also worth further exploration. Team members used words such as “trustworthiness,” “integrity,” “openness,” “professionalism,” “academic bounds,” and “honesty” when describing the values of the team they expect a central database to represent. These expectations show that a central database is viewed amongst these archaeologists as having value to the project as more than a way to organize data, but as a symbol to those outside of the project that represents the team’s character and commitment.

Team members' expectations of a central database to impact the team's dynamic are also worth exploring, especially when considering the possible sociotechnical aspects of databases. The expectations for the database to allow transparency among team members shows the team's view of a central database as a tool for communication. Hine (2006) specifically discusses the extent to which information and communication technologies such as databases provide new communication regimes and forms of collaboration. Hine (2006) cites Hilgartner's 1995 discussion on communication regimes which are "established networks for the dissemination of science, comprising actively constituted systems of technical and social arrangements". Hine (2006) notes the thoughts further discussed about these communication regimes are not well suited for thinking about the "small-scale uses" of databases where they act as "research tools or instruments" which was the focus of her paper as well as this case study. She mentions journals as the iconic regime, but notes Hilgartner's suggestions on how databases allow for the development of new regimes, and how each regime he discusses "involves a database but puts in place different arrangements for acquiring, validating, and circulating data, for determining the extent to which the system is centralized, and for establishing the relationship with journals". For the small-scale use of a future database in this team's workflow, thinking about the extent to which this system will be centralized is necessary to move forward, as well as thinking about the relationship of a future database not to journals, but perhaps articles published by the team, and reports (as was mentioned as an expectation to shorten written reports by providing a reference to object information in a central database).

Several team members desired the ability to use the database as a way to understand what is happening in other areas of the project, as a way to view their work in the context of others' work, and as a means to apply the data of others to help solve questions. Faniel et al. (2018) also mentions context and encourages specialists to use the data of other team members so their studies will be "more contextually informed". In regard to using one system, they state that combining the data of each member also has implications for data reusers by enabling "better evaluation of data." This relationship between databases and evaluation was mentioned briefly in the literature review. For example, Hine (2006) discussed issues raised by databases such as changes to the way research is evaluated and the work practices of scientists. From several team members' expectations in this study the idea of using others' data so one's own work is more contextually informed is clearly seen through examples such as the need for context in determining conservation treatment.

The isolation that some team members expressed feeling is also described by Faniel et al. (2018) as "specialist silos" where the datasets of specialists are "siloed bodies of inconsistently managed data and documentation" (p.114). Some of their suggestions for

moving out of the silos include 1) developing a policy that specialists agree to before participating in the project that outlines expectations for data sharing, analysis timelines, and publication, 2) facilitating discussion on how specialist data can be integrated into the “primary excavation dataset”, and 3) providing the director with work summaries and drafts of data from each specialist before they leave the site (p.114).This last suggestion is one the team already implements, however, as one member mentioned, this does not always happen while on site due to time constraints. Can a central database help solve this? With a central database, several team members noted they expect the director will have access to everyone’s data without having to retrieve files from various members. Though having a central system will likely decrease the time it takes to access information, there is still an issue of the time it takes to input data. Some team members realized that even with a central database, if members are not entering their data into the system in a timely fashion, problems dealing with access to information can still occur. To avoid these problems several members noted the team will need to specify designated time in the field solely for data entry in order for a central database to be effective.

A relational database solves the roots of some of the team’s higher level problems such as the time it takes to get information. Like one member explained, the primary issue is not having access to information when one needs it. A central system where all of the sub-teams enter their records so they are linked allows everyone to see how the material they work with is connected to other areas of the project, and allows team members to gather information without needing to first contact other members, wait for that person to find the information, and then wait for that person to respond.

Conclusion

From this case study we find that some archaeologists do not simply view a central database as a place to store and retrieve data, but as a tool that preserves the legacy of its team members and allows for transparency amongst members and the Egyptian antiquities authorities. Such a tool is expected to reduce confusion of finding information among different systems and different people by having all of the information in one place. As a result, this team hopes that having their data centralized will decrease the time spent trying to find and access information and allow for faster publication.

Understanding the motivations and expectations for creating central database systems for archaeological projects has implications for archaeologists, information scientists, and information technology specialists. This study dives into the reasons why implementing a central database is desired for archaeological work and provides an example of how an archaeology team thinks about databases. Information scientists have the capability to serve as a bridge between IT and archaeology through user research and gaining an understanding of data curation practices in archaeology. IT professionals benefit from this research by applying user research to the systems they create, which directly benefits archaeologists' capabilities to store, track, and communicate data with one another. In the future I would like to continue work on understanding how archaeologists view database systems, specifically the sociotechnical aspects, and further explore the ideas of databases as a representation of team values and academic legacy.

References

Akmon, D., Zimmerman, A., Daniels, M., & Hedstrom, M. (2011). The application of archival concepts to a data-intensive environment: working with scientists to understand data management and preservation needs. Archival Science, 11(3), 329–348. https://doi.org/10.1007/s10502-011-9151-4

Bietz, M. J., & Lee, C. P. (2009). Collaboration in Metagenomics: Sequence Databases and the Organization of Scientific Work. In I. Wagner, H. Tellioğlu, E. Balka, C. Simone, & L. Ciolfi (Eds.), ECSCW 2009 (pp. 243–262). London: Springer London. https://doi.org/10.1007/978-1-84882-854-4_15

Bria, R., & DeTore, K. (2016). 1.5. Enhancing Archaeological Data Collection and Student Learning with a Mobile Relational Database. Mobilizing the Past. Retrieved from https://dc.uwm.edu/arthist_mobilizingthepast/7

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project. Field Methods, 23(2), 136–155. https://doi.org/10.1177/1525822X10388468

Faniel, I., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J., & Yakel, E. (2013). The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse. Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries, 295–304. https://doi.org/10.1145/2467696.2467712

Fearon, D. S., Jr., Borgman, C. L., Traweek, S., & Wynholds, L. (2010). Curators to the Stars. Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47, 162:1–162:2. Retrieved from http://dl.acm.org/citation.cfm?id=1920331.1920534

Federici, T. and Braccini, A. (2012). The Interplay between Practitioners and Technological Experts in the Design Process of an Archaeology Information System. Journal of Cases on Information Technology, 14(1), pp.26-45.

Gidding, A., Levy, T., and DeFanti, T. (2014). ArchaeoSTOR: The Development and Utilization of a Web-Based Database for the Field and Lab. Near Eastern Archaeology, 77(3), p.198.

Hein, A. and Kilikoglou, V. (2011). ceraDAT-Prototype of a Web-based Relational Database for Archaeological Ceramics. Archaeometry, 54(2), pp.230-243.

Hine, C. (2006). Databases as Scientific Instruments and Their Role in the Ordering of Scientific Work. Social Studies of Science, 36(2), 269–298. https://doi.org/10.1177/0306312706054047

Jamieson, A. (2015). Developing Strategies for Sustainably Managing Archaeological Collections. Journal of Eastern Mediterranean Archaeology and Heritage Studies 3(1), 71-77.

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

Kansa, S. W., Kansa, E. C., & Schultz, J. M. (2007). An Open Context for "Near Eastern Archaeology." Near Eastern Archaeology, 70(4), 188–194. https://doi.org/10.2307/20361331

Kintigh, K. (2006). The Promise and Challenge of Archaeological Data Integration. American Antiquity, 71(3), 567–578. https://doi.org/10.2307/40035365

Labrador, A. (2012). Ontologies of the Future and Interfaces for All: Archaeological Databases for the Twenty-First Century. Archaeologies, 8(3), pp.236-249.

La Bella, L., & Nagle, J. (2014). *How Do I Use a Database?* Retrieved from http://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=1779125

Marwick, B. (2016). Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. Journal of Archaeological Method and Theory, 24(2), pp.424-450.

Matthew L. Vincent, Falko Kuester, & Thomas E. Levy. (2014). OpenDig: Digital Field Archeology, Curation, Publication, and Dissemination. Near Eastern Archaeology, 77(3), 204-208. doi:10.5615/neareastarch.77.3.0204

Palmer, C., Weber, N. M., Renear, A. H., & Muñoz, T. (2013). Foundations of Data Curation: The Pedagogy and Practice of "Purposeful Work" with Research Data. Retrieved from https://www.ideals.illinois.edu/handle/2142/78099

Palmer, C. L., Thomer, A. K., Baker, K. S., Wickett, K. M., Hendrix, C. L., Rodman, A., … Fouke, B. W. (2017). Site-based data curation based on hot spring geobiology. PLoS One; San Francisco, 12(3), e0172090. http://dx.doi.org.proxy.lib.umich.edu/10.1371/journal.pone.0172090

Poehler, E. (2016). 1.7. Digital Pompeii: Dissolving the Fieldwork-Library Research Divide. Mobilizing the Past. Retrieved from https://dc.uwm.edu/arthist_mobilizingthepast/9

Sagrario Resurreccion Simbulan, M. (2013). Transitioning from Data Storage to Data Curation: The Challenges Facing an Archaeological Institution. Issues in Informing Science and Information Technology, 10, pp.489 - 499.

Schloen, J. D. 2001. Archaeological Data Models and Web Publication Using XML. Computers and the Humanities 35(2): 123–52.

Smith, N.G., Karasik, A., Narayanan, T. et al. J Archaeol Method Theory (2014) The Pottery Informatics Query Database: A New Method for Mathematic and Quantitative Analyses of Large Regional Ceramic Datasets 21: 212. https://doi.org/10.1007/s10816-012-9148-1

Von Ofen, U. (1993). Relational Database — Structures in Archaeology ADS — an Application in Client-Server-Conception Developed by Means of CASE Method. Historical Social Research / Historische Sozialforschung, 18(3 (67)), 22-34. Retrieved from http://www.jstor.org.smithsonian.idm.oclc.org/stable/20755764

Yakel, E. (2007). Digital curation. OCLC Systems & Services: International Digital Library Perspectives, 23(4), 335–340. https://doi.org/10.1108/10650750710831466

Appendix

Content List:

1. Interview Protocol

Appendices 1

Interview Protocol

My name is Jasmine Smith and I am a master's student at the University of Michigan School of Information. I'm interested in studying Egyptology in the future and have a background in archaeology, so I decided to do a thesis to figure out what are the intersections of Egyptian archaeology and information science. Luckily, I talked to [director's name] who told me about this database project, which was a perfect fit. For my thesis I'm trying to answer 2 main questions:

- What are the motivations for a central database?
- How will a central database affect data collection and dissemination of information about the site?

From this exchange I hope to better understand:

- What's not working with the team having multiple databases or other ways of storing data?
- How the team thinks a central database will improve this
- What are the effects or impacts of having multiple databases on the work and scholarship you're producing?

I could not be more pleased that you have set some of your time aside for us to conduct this interview. This interview will take around 60 minutes, during which I will be asking you questions about your role on the team, your preferences for data collection and management, and your opinion on the purpose for building a central database. I ask that you please treat me as someone not familiar with the practices of Egyptian archaeology. You are the expert on this information, and I am here to learn from you. Before we start, I want to make clear that if at any point you become uncomfortable answering any of the questions, please feel free to let me know and I will either move on to a different field of questions or conclude the interview. Again, I appreciate your willingness to meet with me, and recognize this interview is voluntary on your part. I can end the interview for any reason with no repercussions, and even destroy the notes taken at your convenience. Furthermore, to the extent possible, all of your comments and contributions will be confidential. The information gathered from the interviews will be compiled, presented, and viewed in aggregate, with no distinguishing information about which department or individual raised certain concerns. If there are any particular statements, or responses you would like me to keep completely off the record please feel free to let me know and I will adjust my transcripts accordingly. Before we proceed, would you mind if I took an audio recording of the remainder of this interview just to ensure I don't miss anything? No one other than I will have access to this recording at any point. If you are uncomfortable with this, I completely understand and would be happy to proceed without it. Do you have any questions or concerns for me? Excellent, let's begin.

- Introduction
  - Can you tell me about your role at the site?
    - How long have you been working there?

            - How long have you been working in your area of expertise?
            - What are your research questions or goals, why is the study of the site important?
    - Data Collection and Management
        - Walk me through your process of working with an object, painting, unit, etc. and what information you are recording
        - Can you tell me about what you're using for a database or how you store your data now and how you use data?
            - Why did you start using your current system of storing data?
            - What do you like and dislike about it?
        - How has data management changed at the site since you started working there?

    - Overarching
        - What about your work/ your data / the team's data collection in general do IT experts need to know in order to make something that will fit the needs of the project and not deliver a cookie cutter product?
        - What about the site needs to be represented in the database (in terms of your work and overall?)
    - Purposes of creating the database
        - What went into the decision to create a central database?
        - What would you say are the 3 main motivators to creating it?
        - How do you expect this database to affect your work and work at the site in general?
            - Do you expect this database to affect data collection and curation at the site?
        - In your opinion, what are the end goals of having a centralized database?
        - How do you think having a central database will help the team achieve their research goals and mission?
    - Specifics about the database and how you envision it being used
        - Based on your work at the site, what would be an example of an 'entity' that would be described in the database? (e.g. a statue, bowl, etc)
            - What attributes of this entity would be entered into the database to describe that entity? (e.g. what would the fields in the database be? What information about that entity needs to be in the database?
        - Will data be entered directly from the site, or entered in later
        - What other entities would you include in the database in terms of what you work with at the site?
        - Who will be responsible for entering data into the database?
        - Who will be the users of this database(researchers, just the field team, etc)
        - How should the data be backed up, will it be available through the cloud?

        - How will it be used for retrieval?
            - Do you need to export data or reports?
            - Does it need to be exported in a certain file format?

- Issues
    - Do you have past experience making databases or bringing different data together? Has the team tried to work with IT specialists or database administrators?
    - Are there any issues or concerns you foresee with implementing a central database into the team's workflow?
        - Any technical issues?

- Conclusion
    - What would you consider a big success for this database project going forward?
    - Is there anything else I didn't ask about that you think is important for understanding your role in data management at the site?