2016

Cleaning up Minnesota's Archeological Record with MAID: The Minnesota Archeological Integrated Database

Andrew Allen Brown
Minnesota State University Mankato

Cleaning up Minnesota's Archeological Record with MAID: The Minnesota Archeological Integrated Database

By

Andrew Allen Brown

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

In

Anthropology

Minnesota State University, Mankato

Mankato, Minnesota

July 2016

7/7/2016

Cleaning up Minnesota's Archeological Record with MAID:
The Minnesota Archeological Integrated Database

Andrew Allen Brown

This thesis has been examined and approved by the following members of the student's committee.

Dr. Ronald C. Schirmer, Advisor

Dr. Edward P. Fleming, Committee Member

Dr. Phillip H. Larson, Committee Member

Abstract

Minnesota archeologists face many difficulties in conducting archeological research and managing the state's cultural resources such as a lack of standardized data formats and field/lab procedures, a lack of a centralized data repository, and insufficient existing databases. The purpose of this thesis is to build the foundation for a database system that addresses these difficulties along with being efficient and effective for entering, managing, and analyzing archeological data produced in the field and in the lab. The Minnesota Archeological Integrated Database is being built to be a long-lasting, constantly evolving system to be used by archeologists and cultural resource managers for years to come.

Contents

1. Introduction

More than 18,000 archeological sites have been recorded (Minnesota Office of the State Archaeologist 2015, p.11) in Minnesota but the only database to contain records for each is housed at the Minnesota Office of the State Archeologist (hereinafter OSA; Minnesota Office of the State Archaeologist 2014). The records themselves contain only a fraction of the existing information for each site (see section 2.1) and does not include complete artifact catalogs, provenience information, photographs, or supplemental contextual data (e.g., soil data) that are often housed elsewhere, yet which are necessary for understanding a site and making correct inferences about it, and decisions about its characteristics and importance. Therefore, archeological data in Minnesota is quite difficult to utilize–preventing cultural resource managers from effectively preserving Minnesota's archeology and research archeologists from efficiently conducting research. Provenience information is of considerable significance as the interpretation of archeological remains relies heavily on context (Butzer 1980). Many records have not been digitized and those that have been digitized are often in excel spreadsheets with project-specific data structures that may prevent effectively combining datasets for analysis due to differences in data structure or format (section 2.1).

The purpose of this thesis is to build the foundation for a database system that addresses the aforementioned difficulties in archeological work in Minnesota, along with being efficient and effective for entering, managing, and analyzing archeological data (artifactual and provenience) produced in the field and in the lab. The Minnesota Archeological Integrated Database (hereinafter MAID) is being built to be a long-lasting, constantly evolving system to be used by archeologists and cultural resource managers for years to come. The bulk of the thesis is the working system itself and the current document is meant only to provide background information for its need and the philosophy guiding its development. The majority of the system's documentation will be available online and updated as the system develops.

2. Problems Necessitating MAID's Development

2.1 Lack of Centralized Digital Data Repository

The lack of a centralized archeological data warehouse, digital or otherwise, has made building a truly synthetic presentation of the prehistory of Minnesota nearly impossible. OSA has traditionally been responsible for generating and assigning official state site numbers to archeological sites but duplicate sets of official files were created at MHS Archaeology Department at Fort Snelling, the University of Minnesota, and Hamline during the 1970s (Koenen 2005). The Minnesota State Historic Preservation Office (SHPO) and various individuals throughout the state were creating their own site forms and using their own numbering systems in the 1980s (ibid.). Therefore, original site documentation must be sought in numerous places besides the OSA; OSA tracks burial data while repatriation status is maintained by the Minnesota Indian Affairs Council (MIAC), and SHPO tracks National Register status. The majority of archeological project data, including catalog sheets, field notes, photographs, etc., is usually stored at the company, agency, or institution that originally conducted the fieldwork, creating a deep break in the continuity of the archeological record. Researchers interested in analyzing collections are not the only ones suffering due to this lack of organization. Cultural Resource Managers (hereinafter CRMs) are unable to properly perform due diligence without a system for integrating relevant information such as locations surveyed during, and artifactual content of past projects. This situation creates an unacceptable risk to the archeological record itself.

The Minnesota Archaeological Inventory Database (Minnesota Office of the State Archaeologist 2014), as mentioned above, is the only comprehensive archeological database in the State. OSA (2015, p.10) presented "four common sources of error" associated with the site database:

> 1) the original data reported on the site form may be inaccurate;
> 2) the data reported on the site form may be a unique interpretation or have inconsistent interpretations by different archaeological investigators;
> 3) correct data from a site form may have been incorrectly entered into the database;
> 4) different data input personnel prior to 2007 may have used inconsistent codes for the data.

Researchers and CRMs may be required to contact several institutions, including the Minnesota Historical Society (hereinafter MHS), OSA, MnDOT, U.S. Army Corps of Engineers, National Park Service, Department of Agriculture, Department of Natural Resources, etc., in order to determine the extent to which data have been collected from a site or project area. Afterwards it may be that the records are too sparse (general lack of data[1]), poorly documented (missing data[2]), outdated, or erroneous (typos[3]). At times it is not even clear what institutions to contact.

[1] e.g., a site location was documented but no further information was recorded.
[2] e.g., a site was recorded and artifacts were cataloged but the location of the catalogs are not included on the site form.
[3] e.g., site records in the OSA database often contains typos

2.2 Lack of Standardized Cataloging Format

The data produced through archeological work in the state are scattered among the various institutions in the region that manage cultural resources and conduct data collection—Minnesota State University, Mankato, MHS, US Forest Service, US Army Corps of Engineers, Science Museum of Minnesota, numerous other colleges and universities, Tribal Historic Preservations Offices, etc. Each institution has implemented their own protocols for recording data both in the field and in the lab, including different cataloging software (EMu, Re:Discovery Proficio, PastPerfect, etc., see section 2.4) and field data collection forms. The differences between these protocols are such that researchers cannot simply combine datasets for analysis because of the variation in artifact classification and basic data format. Additionally, terminologies used, for example, to describe lithic raw material and tool types, change over time and the small differences between the institutions are exacerbated to the point that whole collections often need to be re-analyzed. At a minimum, the data cannot be accepted at face value. Furthermore, neither OSA nor MHS are able to produce a verified, comprehensive list of all repositories in the state, not to mention what collections they have, where they came from, and what they contain (Bruce Koenen, personal communication, 2012; Patricia Emerson, personal communication, 2012).

2.3 Unstructured and Inconsistent Provenience Data

A complete record of previous archeological surveys and excavations is essential for preventing areas from being unnecessarily surveyed multiple times and identifying areas to target for future gap-filling research. It is becoming more common to record general survey areas and specific provenience units with high-precision GPS units but, like artifact cataloging, there are differences between the terminologies used to record—e.g., excavation units vs pits—and the structure used to store the data—e.g., flat tables vs relational database systems. The varying ways provenience is recorded, stored, and shared has led to an obfuscated record of exactly what work was conducted and where it occurred. Instead of having ready access to this information, it has become necessary to research past research in order to actually conduct research. This may be inevitable but a better system for storing and querying previously conducted archeological projects will lessen the amount of work needed to plan future work.

2.4 Current Database Solutions Insufficient

Three popular database systems designed with archeologists in mind were considered during MAID's development: Proficio, EMu, and PastPerfect. The purpose of this section is not to present a thorough discussion of the features provided by each software option but to provide reasons why each is insufficient for the needs of the Minnesota State University, Mankato Museum of Anthropology (MSU). The main issues with existing database solutions that are addressed in this section are price, flexibility, and system performance and longevity. For a comprehensive discussion of the software discussed in this section, see Thomson 2014 and Miller 2012.

Curation databases are designed and used primarily for tracking the ownership, location, and condition information of objects, and are typically the choice of repositories that may or may not have archeological practitioners on staff (Thomson 2014). Research databases, on the other hand, tend to be focused on an artifact's provenience information and taxonomic classification for the purposes of conducting analysis, and are most often used by practitioners who are not specialists in curation. Research databases are often as simple as a couple spreadsheets or a Microsoft Access database and are often highly-tailored to a project (e.g. thesis or dissertation) or site for which a specific and detailed report is required. Given the primary users (repository staff), many archeological database systems in use were designed with curation rather than research in mind and, while these systems allow curators to manage their collections, they are poorly designed for research purposes. Systems capable of handling the complexity of archeological research tend to be expensive[4], limiting the degree to which small organizations can participate in the archeological process. Price may not be a significant issue for larger institutions (e.g., Science Museum of Minnesota, MHS, etc.) but smaller institutions (e.g., county historical societies and local museums) may be unable to afford the licenses and training required to use expensive software.

The constantly changing technology industry means that inevitably software, including databases, will become obsolete. PastPerfect, for example, was developed using Microsoft Visual FoxPro, a programming language no longer being supported by Microsoft (PastPerfect 2015). While PastPerfect is no longer being developed using FoxPro, the loss of Microsoft's support demonstrates the importance that continuous development has on the lifespan of a database system (PastPerfect 2015).

Commercial database systems designed for generic collections management tend to have limited support for deeply nested hierarchical data. Proficio provides the ability to create and manage provenience units and hierarchically structure them in the database.

[4] EMu costs ~ $17,000/5 users, Proficio costs ~$1300/2 users (Schmitt 2014)

However, as of version 8.18, provenience units cannot be nested deeper than 2 levels, which makes it difficult to structure complex provenience units as we may encounter in the field. A common example of hierarchical provenience is the grouping of excavation units together into a block, which may only relate to proximity but often has to do with a shared grid datum providing each member of the block a grid reference from which all measurements are based. Each excavation unit has one or more vertical levels that may or may not be further subdivided. Provenience units are not always mapped in this cleanly organized hierarchical manner as feature halves and quarters are often mapped relative to the feature center instead of an excavation unit or block corner which, in turn, are mapped relative the site datum. Relative datum points may vary at each level of a hierarchical provenience system. This type of nested provenience system is fairly common in archeological field practice and requires that a database system be able to handle such an example intuitively and easily. Proficio users are able to define “master contexts”, which represent groups of provenience, and “contexts”, which represent individual provenience units or provenience subdivisions (e.g., vertical levels). Contexts cannot be nested within other contexts, limiting how closely provenience in the database mirrors provenience as defined in the field.

The price of archeological and curation database systems is a major issue for smaller institutions like local museums and county historical societies as they are often run by volunteers and receive little funding. For this reason, it is important for database software to provide the capabilities outlined above as well as be affordable enough for small institutions. EMu was excluded as a potential software package for this reason (see footnote 1 above for price) as price is a major issue for smaller institutions.

2.5 Outdated Historic Contexts

The Secretary of the Interior's Standards mandated the development and maintenance of historic contexts used to guide research in the state through the Resource Preservation Planning Process (Secretary of the Interior 1983). Minnesota's contexts were compiled in the mid-1980s (Dobbs et al. 1988). Context development used existing knowledge at the time, including the types and distributions of sites, site types, and artifact types to create study units that are intended to guide interpretation and identify gaps in knowledge, thus they are important in determining what work needs to be done. Maintenance and updating was an explicit requirement of the Standards, but as described above, fragmentation precludes this from happening. In turn, this exacerbates other problems because some data at the OSA and SHPO, and many cataloguing systems, refer to the existing state contexts such as "Blue Earth Oneota" or "Silvernale", and information like artifact type and style names (e.g. Anfinson et al. 1979) like "Onamia", that are now obsolete. This creates a self-sustaining source of error because new data are not fed back into the study context, which should be automatically reevaluated as artifacts are entered into the system.

2.6 Lack of Supplementary Contextual Data

Archeologists require complete and thorough data in order to conduct research. These data often include much more than curation information alone. Contextual data include environmental, geomorphological, and soil records, among other things (Taylor 1948; Willey & Phillips 1958; Caldwell 1959; Butzer 1980; Trigger 1991). Accurate provenience information is essential as artifacts cannot be considered isolated from their context. Highly precise provenience is needed for the analysis and interpretation of intra-site features and the cross-comparison of the artifactual content of provenience units. Recording data so that artifacts can be queried spatially, by provenience unit, or type of feature is exceedingly difficult in a flat, tabular format. Provenience, in general, is not often recorded separately in curation-oriented databases as excavation units are not usually curated. While provenience units are not discrete, curated objects like artifacts, they should be treated as carefully because of their contextual significance—artifacts are recovered from these arbitrary spatial boundaries that are later used in analysis to compare artifact assemblages. Archeologists often need to view the wider context of the entire provenience unit in order to interpret its contents, as much of an artifact's interpretive potential lies in its spatial relationships with surrounding artifacts, soils, etc. (Fagan 2011; Price 2006).

2.7 Inefficient Cultural Resource Management

CRM archeology shares many of the same requirements of research archeology, but with a larger focus on site-level interpretation. Managers require the ability to quickly determine potential site conditions in order estimate the type of work needed. Doing this involves using records of pre-contact environmental conditions and post-contact cultural disturbances. A basic summary of the types of archeological materials a survey may yield, if conducted, is also necessary for planning work to establish the required level of effort and to estimate reasonable pricing. In addition to the common archeological necessity of understanding past research, CRM is federally mandated and agencies with land management responsibilities such as the Department of Natural Resources and

Department of Transportation are unable to make fully informed decisions regarding land use practices. These agencies may unknowingly damage archeological sites without a system in place for tracking past archeological work.

3. Building MAID

3.1 Open Source Software

MAID follows the open source philosophy that software performs best through the collaborative and transparent efforts of many individuals (The Open Source Initiative 2007). This philosophy dictates why many of the libraries and software packages used in MAID's development are open source. This is important as it allows anyone interested to become involved and contribute their time and effort to making the software better. Open source does not mean, however, that anyone can simply make changes to MAID's codebase. Only approved individuals have permission to make changes to the code, but anyone can view the code and request changes be made.

3.2 Backend

The backend[5] consists of a web server built with Node.js, a JavaScript runtime that can be used for building server and desktop applications, using the ExpressJS library (Node Foundation 2016b; Node Foundation 2016a). Node was chosen mostly because it enables MAID's developers to write code for the backend and frontend in JavaScript instead of multiple different languages. This makes it easier for new contributors who may not be experienced programmers to better understand the system and how they can contribute.

[5] The part of the system dedicated to managing data and data access.

3.3 Database

The database used is RethinkDB, a fully open source, highly scalable database built "for the real-time web" (RethinkDB 2016). This database was chosen because it offers the flexibility difficult to achieve with a traditional database (see Section 3.3.1) and because it is designed for real-time applications where many users can be interacting with the same data at the same time, e.g. collaboratively editing documents and site records (Section 3.3.2). RethinkDB also offers the ability to store and query geospatial data, which is essential for MAID as nearly all data involved have some sort of geospatial component.

3.3.1 Relational Databases vs. Document-Oriented Databases

Traditional database systems are built using relational databases (Oracle MySQL, PostgreSQL, Microsoft Access, etc.) which store data in tables with rows and columns that represent sets of objects (rows) with defined properties (columns) that can be related to objects in other tables by referencing their unique ID (Oracle 2015). Because of the tabular structure, many tables may be required in order to represent complex objects such as archeological artifacts in a relational database. Therefore, a single artifact's properties may be stored in several tables, not including all the tables required for storing provenience and curation information.

Document-oriented databases (e.g., RethinkDB) store objects using a concept aptly named the "document"—a data structure that does not force the same template/fields on each item, allowing for flexibility in what properties are associated with each object, a primary reason MAID is built with this type of database. Documents can relate to or be embedded within other documents in order to create relationships between object types.

The flexibility that document-oriented databases provides allows for rapid development of data-driven applications.

3.3.2 Realtime-Oriented

Realtime web applications are applications that push and receive notifications of various changes. Two familiar examples of realtime applications are instant messaging apps and collaborative document editors (e.g., Google Docs). Upon user input in a collaborative editor, notifications are sent to other users in "realtime" and the input value is reflected and everyone sees the change.

There are many frameworks that offer realtime capabilities (PubNub n.d.; Firebase n.d.) but these systems work by frequently polling the database for changes, while RethinkDB is able to push changes directly to an application (Akhmechet 2015). This limits the amount of database queries required and provides a boost in system performance and developer overhead.

3.3.3 Database Design

The database was designed to be as flexible as possible to match the inherent uncertainties and ambiguities of archeological data (as discussed in section 2). One issue in cataloging archeological artifacts is that recording errors and mistakes are often made, which can mean, for example, that we no longer know which level of an excavation unit that an artifact came from, but only that the artifact was recovered from an excavation unit. MAID accounts for this by structuring provenience units hierarchically so if an artifact's exact provenience cannot be established a user can relate it to the most specific provenience possible. In the same way, artifact taxonomic classifiers are also hierarchically structured so a user can be as specific as "Banded Prairie du Chien Chert" or as general as “Rock” or “Chert”. This makes it much simpler to import legacy datasets that often contained older, less specific terminology.

The terms and data structure used in MAID are the ongoing byproduct of more than five years of meetings and conversations with practicing archeologists, lab and field personnel, and a steering committee consisting of representatives from MIAC, OSA, SHPO, MHS, MnDOT, and MSU. The interactions and subsequent feedback have provided many useful ideas from both the research and CRM perspectives that have greatly contributed to MAID and this transparent and collaborative process will continue throughout the system's development and use-life.

As this thesis only represents MAID’s foundation, the database schema only contains the essential entity types and relationships needed to model archeological data. See Appendix I for a graphical representation of the database design.

3.4 Interface

MAID’s interface had to address two main issues: availability and interactivity. Availability, in this context, relates to the ease at which users are able to access the system. This could mean viewing an artifact record, looking at images of excavations, checking the status of a collection, or cataloging artifacts. Interactivity relates to the degree to which the interface responds to user action (e.g., the data entry forms update based on user input).

3.4.1 Availability

The simplest option for a multi-institutional database system is for it to be browser-based. An interface used through browsers like Google Chrome or Mozilla Firefox allows as many people as possible to use it with no installation needed. This is important because people may have difficulty installing software on company or government PCs and a web-based system ensures all users are running the same software version.

3.4.2 Interactivity

Archeological and geospatial data entry is very complicated and can be made more efficient and user-friendly with an easy-to-use interface. Javascript is by and large the most dominant language for writing web applications in the browser. The interface is built with React (Facebook 2016), a Javascript library developed at Facebook. This library has become one of the most prevalent tools for building web interface in the last few years; it is the sixth most popular repository on GitHub, with nearly 45,000 stars as of July 2016 (GitHub 2016).

4. Discussion

The benefits MAID provides archeologists are numerous. Perhaps most significantly, MAID is a web-based system useable anywhere with an internet connection (section 4.1). Its provenience-driven design provides artifacts with a higher level of contextual detail (section 4.2). Import/Export flexibility (section 4.3) and a single cataloging system (section 4.4) ease repositories' transitions to MAID from their existing systems. In addition, several projects arose out of MAID and will provide benefits during the artifact cataloging process, the visualization of artifacts, and the interpretation of archeological sites (section 4.5).

4.1 Web-based Interface

The first benefits are the same as any web-based data management system. It is available anywhere an internet connection is found, eliminating the need to request copies of spreadsheets containing data from multiple places to aggregate data. In so doing it also provides a continuously updated data source so that people do not need to rely on outdated information.

4.2 Provenience-Driven Design

MAID's emphasis on provenience provides the potential for an all-in-one archeological data entry system. Provenience units can be created while planning fieldwork or while in the field during excavations/survey. Subsequently, in the lab, artifacts can be easily associated with the correct provenience. Associated documentation that would traditionally be recorded on paper field forms (e.g. excavation unit forms, field specimen logs, photo logs) can be associated with their respective units in the field rather than after the fact, reducing the amount of data reorganization needed after fieldwork is completed. Artifacts can be queried by provenience unit, across provenience units, depths, strata, etc.

4.3 Fluid Data Import/Export Process

Importing data into MAID can be done by entering or reformatting data into a template spreadsheet and loading it into the web interface. A table is populated with the imported data and each entry is validated and errors are flagged and can be edited before saving to the database. Taxon terms in relationship fields are queried and must be valid before saving. This simple process creates a fast and reliable way to incorporate the vast body of data that exist across the state, but which are currently either unavailable or unusable.

Data can be exported based on user-defined queries out of MAID into comma-separated value (CSV), Microsoft Excel, and Javascript Object Notation (JSON) formats. In the future, other formats such as Microsoft Access and ArcGIS geodatabase will be available.

4.4 Standardized Cataloging Format

MAID provides the foundation for a standardized method of collecting and recording data in both the field and the lab. Projects, provenience units, and artifacts are tightly related in the system as they are in practice. Provenience units are created and associated with a project in the system while in the field and artifacts are associated directly with them in the lab. This results in a much lower chance of erroneously recording an artifact's provenience. Recording all data, starting with projects and ending with artifacts, will also make planning future fieldwork and research much easier as everything related to archeological work is found in one place. A significant benefit is that MAID will allow repositories (e.g., MSU, local museums and county historical societies) and agencies (e.g., MnDOT) to meet their federal mandates with regard to data management and availability by providing a single source and format for storing and accessing archeological data in the state of Minnesota.

4.5 Ancillary Projects

There are several ancillary projects that will supplement MAID and are not detailed in this thesis. These include 3D artifact models, historic aerial photography, and a catalog guide.

A methodology and guide for the creation of 3D models of artifacts is being developed and tested. In the future, MAID users will be able to view 3D models as well as photographs from the web interface.

4.5.1 Historic Aerial Photography

The Farm Service Agency aerial photographs (ca. 1938) have been scanned and are being mosaicked and georeferenced for use in MAID. Mosaicking is being conducted using

Agisoft Photoscan (Agisoft LLC 2016). A methodology and example code will be released in the future.

4.5.2 Catalog Guide and Local Taxonomy

An artifact cataloging guide has been written and will be used as the basis for the help documentation on MAID's web interface. Included in the catalog guide is a taxonomy designed specifically for Minnesota archeology that can be expanded to include other regions. There has been great interest in the guide and taxonomy from researchers in Minnesota as Minnesota lacks a standardized taxonomy used in archeological cataloging. The Getty Art and Architecture Thesaurus (The Getty Research Institute 2016) used by the largest repository in the state (MHS) does not contain many entry options that are frequently required by prehistoric archeologists in Minnesota. The catalog guide and associated taxonomy will be a benefit to Minnesota archeologists by including locally relevant terms (e.g., lithic raw material types, pottery types, etc.) not found in the Getty AAT. In the future, the taxonomy will be a large part of MAID and experts in the state will be able to continuously add new terms and revise existing terms.

4.5.3 3D Artifact Models

The MSU recataloging process during MAID’s development included photographing every diagnostic artifact cataloged, and a methodology is being developed for the rapid creation of accurate 3D models of each. These models will provide users a more complete look at the artifacts; isolated 2D pictures and written descriptions are seldom capable of detailing complex pottery decorations, in particular.

5. Conclusions and Future Work

There are many aspects of MAID that still need to be developed. The most important features planned are a mapping interface and the ability to add photographs and documents to collections and field records. A record-level permissions system is also being planned that will allow records (artifacts, collections, provenience units, etc.) to be managed individually with only certain people allowed specific read/write abilities (if the record owners so choose).

MAID will be hosted at MSU for the foreseeable future and funded in part by MnDOT and the Federal Highway Administration. A team at MSU, working in collaboration with MnDOT and MnGeo is dedicated to developing, maintaining, and supporting the system. As the tools and software used to build MAID become outdated or better alternatives are discovered they will be implemented without requiring anything of the users.

MAID is also being developed in order to connect to other database systems like the Office of the State Archeologist's official site database. The OSA database, in particular, is currently being developed in cooperation with MnGeo and MnDOT. This will allow OSA to manage the official state site records and automatically request the most up-to-date artifact and provenience data from MAID (and vice versa). This functionality will need to be implemented in the future.

In conclusion, the Minnesota Archeological Integrated Database will provide massive benefits to the archeological community in Minnesota by increasing availability of data and providing a standardized method of recording, editing, and viewing data. The ability to view as much of the information necessary to interpret archeological materials in one place will make both research and CRM much more efficient.

6. References

Agisoft LLC, 2016. Agisoft Photoscan. Available at: http://www.agisoft.com/ [Accessed August 1, 2016].

Akhmechet, S., 2015. Advancing the realtime web - RethinkDB. Available at: https://www.rethinkdb.com/blog/realtime-web/ [Accessed August 12, 2016].

Anfinson, S.F. et al., 1979. A Handbook of Minnesota Prehistoric Ceramics. *Occasional Publications in Minnesota Anthropology*, 5.

Butzer, K.W., 1980. Context in Archaeology: An Alternative Perspective. *Journal of Field Archaeology*, 7(4), pp.417–422.

Caldwell, J.R., 1959. The New American Archeology. *Science*, 129(3345), pp.303–307.

Dobbs, C.A. et al., 1988. *Outline of Historic Contexts for the Prehistoric Period (CA. 12,000 B.P. - A.D. 1700)*,

Facebook, 2016. React. Available at: https://facebook.github.io/react [Accessed July 4, 2016].

Fagan, B., 2011. *Archaeology: A Brief Introduction* 11th ed., Routledge.

Firebase, Firebase. Available at: https://www.firebase.com/ [Accessed August 12, 2016].

GitHub, 2016. Most Starred Repositories. Available at: https://github.com/search?q=stars:%3E1&s=stars&type=Repositories [Accessed July 4, 2016].

Koenen, B., 2005. Minnesota's Archaeological Site Files - Their History and Organization.

Miller, T.M., 2012. *Specify for Archaeology: A Proposed Data Model for Archaeological Collection Database Management*. University of Kansas.

Minnesota Office of the State Archaeologist, 2014. Archaeological Inventory Database, Minnesota. Available at: http://www.mngeo.state.mn.us/chouse/metadata/arch_inv.html [Accessed July 31, 2016].

Minnesota Office of the State Archaeologist, 2015. *Archaeology in Minnesota – 2014 Annual Report of the State Archaeologist*, Available at:

http://mn.gov/admin/assets/2014ArchaeologyAnnualReport_tcm36-187061.pdf.

Node Foundation, 2016a. Express. Available at: http://expressjs.com [Accessed July 4, 2016].

Node Foundation, 2016b. Node.js. Available at: https://nodejs.org/en/ [Accessed August 12, 2016].

Oracle, 2015. A Relational Database Overview. Available at: https://docs.oracle.com/javase/tutorial/jdbc/overview/database.html [Accessed August 12, 2016].

PastPerfect, 2015. PastPerfect Development Update: 11/11/2015. Available at: http://museumsoftware.com/roadmap.html [Accessed August 12, 2016].

Price, D.T., 2006. *Principles of Archaeology* 1st ed., McGraw-Hill Humanities/Social Sciences/Languages. Available at: https://www.amazon.com/Principles-Archaeology-T-Douglas-Price/dp/0073271322.

PubNub, PubNub. Available at: https://www.pubnub.com/ [Accessed August 12, 2016].

RethinkDB, 2016. RethinkDB. Available at: https://www.rethinkdb.com [Accessed July 4, 2016].

Schmitt, B., 2014. Collections Management Systems and Vendors. Available at: http://carlibrary.org/CMS-Table.htm [Accessed August 14, 2016].

Secretary of the Interior, 1983. Archeology and Historic Preservation: Secretary of the Interior's Standards and Guidelines. *Federal Register*, 48(190), pp.44716–44742.

Taylor, W.W., 1948. *A Study of Archeology*,

The Getty Research Institute, 2016. Art & Architecture Thesaurus Online.

The Open Source Initiative, 2007. The Open Source Definition. Available at: https://opensource.org/osd [Accessed August 9, 2016].

Thomson, K., 2014. *Handling the "Curation Crisis": Database Management for Archaeological Collections*. Seton Hall University.

Trigger, B.G., 1991. Distinguished Lecture in Archeology: Constraint and Freedom-A New Synthesis for Archeological Explanation. *American Anthropologist*, 93(3), pp.551–569.

Willey, G.R. & Phillips, P., 1958. *Method and theory in American archaeology*,

7. Appendices

7.1 Appendix 1: Database Schema