

Standard Descriptive Vocabulary and Archaeology Digital Data Collection

Caroline Beebe

*The Naming of Cats is a difficult matter,
It isn't just one of your holiday games;
You may think at first I'm as mad as a hatter
When I tell you, a cat must have THREE
DIFFERENT NAMES.*
— T.S. Eliot, "The Naming of Cats," 1939

The naming of things related to digital archaeology data is a difficult matter rarely explicated by textbooks or repositories. As an information scientist and data manager for the Chau Hiix Project in Belize, I consider the issues of data collection, management,

and archiving through the lens of standardized vocabulary. The ultimate goal of archiving site data for validation through reproducible results or future repurposing is basically a problem of communication grounded in vocabulary.

As excavation begins, the archaeologist provides an analog description of their observations. Digital data may be born in the field or turned digital in the lab or during analysis. Digital files containing values and codes, named and organized into folders, represent the characteristics and survey locations of archaeological objects. Consistency in these alphanumeric representations is essential to data discovery for repurposing or reproducibility. Without standardized vocabulary, the researcher must perform three digital searches to collect data about those baked clay

ABSTRACT

Archaeology has embraced the shift to digital technology for collecting, analyzing, and sharing data. Digital repositories are now recognized as essential for data stewardship and are setting standards for data deposition. These new technologies and systems support the scientific need for reproducible results through intra-cultural as well as cross-cultural hypothesis testing. Methods of digital data collection in the field, however, are often site specific, restricted by the limited availability of digital technologies, or not well suited for creating systems that support the requirements of the new digital information paradigm. As a small science project, the Chau Hiix Project in Belize will provide examples of the pitfalls in and insights about shifting to digital technology to make its primary data shareable and reusable. These experiences suggest the need for an international collaborative agenda that develops digital data description standards based on controlled vocabulary, facet analysis, and crosswalks implemented at the analog point of collection.

La arqueología ha aceptado con entusiasmo la transición hacia la tecnología digital para la recolección, análisis e intercambio de datos. Los repositorios digitales son ahora reconocidos como herramientas esenciales para la administración de datos y están estableciendo normas para el almacenamiento e intercambio de información. Estas nuevas tecnologías y sistemas apoyan la necesidad científica de obtener resultados reproducibles a través de la comprobación de hipótesis tanto intra como interculturalmente. Sin embargo, los métodos de recolección de datos digitales en el campo a menudo son específicos para cierta ubicación, están restringidos por la limitada disponibilidad de tecnologías digitales o enfrentan dificultades para crear sistemas que soporten los requerimientos del nuevo paradigma de información digital. Como un pequeño proyecto de ciencias, el Proyecto Chau Hiix en Belice proporcionará ejemplos de las perspectivas y dificultades relacionados con la transición a la tecnología digital con el objetivo de permitir el intercambio y reutilización de sus datos primarios. Estas experiencias sugieren la necesidad de un programa de colaboración internacional que desarrolle estándares de descripción de datos digitales basados en un vocabulario controlado, el análisis de facetas y la implementación de puentes en los puntos análogos del proceso de recolección de la información.

Advances in Archaeological Practice 5(3), 2017, pp. 250–264
Copyright 2017 © Society for American Archaeology
DOI:10.1017/aap.2017.15

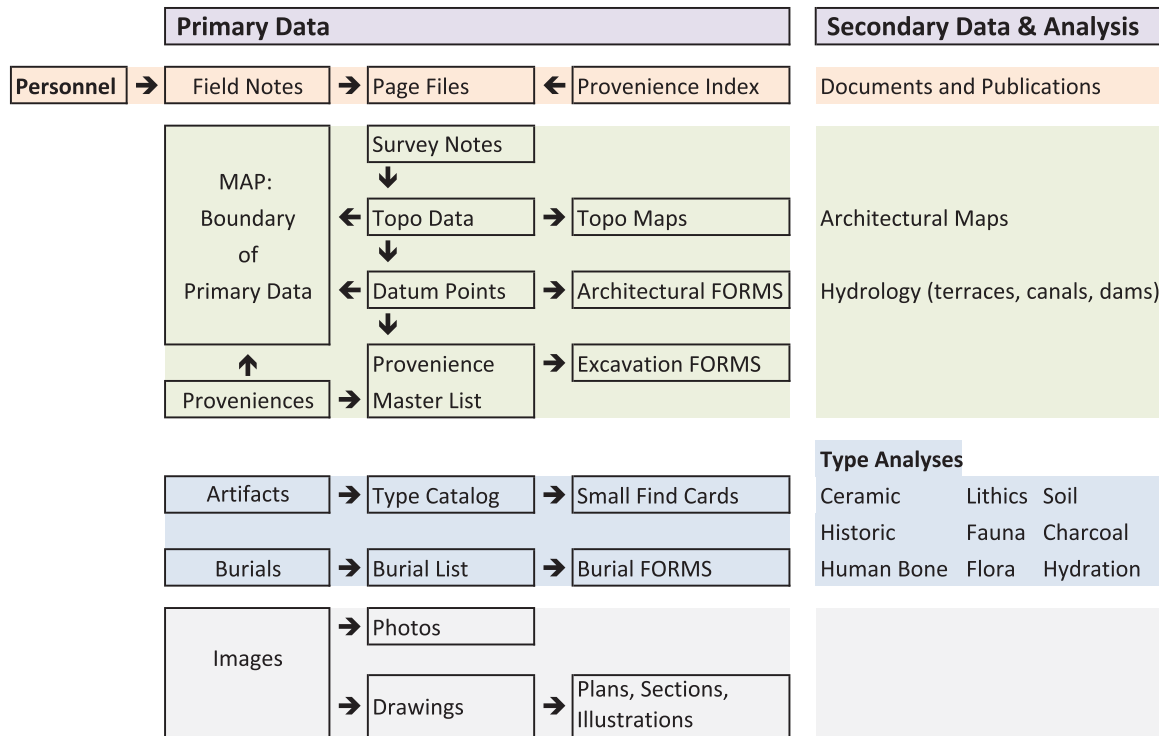


FIGURE 1. Overview of Chau Hiix data collection.

objects referred to independently as ceramics, pots, or sherds. Without consensus on how data are represented in the digital future, data validation and repurposing will be a frustrating endeavor leaving each project as a silo.

Chau Hiix is a small science project that has worked for the last 25 years to develop a comprehensive digital data archive for open publication. Digital technology, however, brings representation problems that were not required in analog recordkeeping. Data repositories tell us what types of files to create but not how to organize and name the data within. How are filenames and object names standardized; how are location schemes constructed? The following description of the Chau Hiix solutions provides an example of a simple, spreadsheet-based data collection plan and how names were created for the objects, locations, and files.

DATA MANAGEMENT PLANNING

Data repositories such as the Archaeology Data Service (ADS) (2011) and the Digital Archaeology Record (tDAR) (Digital Antiquity 2013) advocate that a data management plan be developed at a project's inception and provide best practices and guidelines for data deposition. The repositories guarantee that if you follow their format requirements the data will be migrated and accessible for future scientists who may be using digital analysis tools such as geographical information systems (GIS) or computer-aided design (CAD). Projects incorporate repository format requirements as they operationalize the research design into a data management plan. As archaeology moves into a completely digital environment, how digital data are organized and

represented within those formatted files is the first step: the data collection plan (Figure 1).

The data collection plan describes the words or numbers used to represent the data. This includes the terms used for each material type, how survey codes represent location, how photographs are named to indicate represented objects, how personnel information is managed, and more. The plan also documents how file naming, version control, and backups are handled during data collection and building the archive. The files themselves are formatted according to the repository requirements, guaranteeing access and migration. Data inside the files, however, are discoverable and comparable due to content that reflects standardized terminological representation.

Chau Hiix Digital Data Collection Plan

Chau Hiix is a Maya site located on a seasonal lagoon in the jungles of Belize. The research design focuses on intersite political relations to test competing models of ancient Maya political economy using complementary lines of evidence: area populations, agricultural features, Postclassic settlement, and a long-lived civic center (Andres 2005; Andres and Pyburn 2004; Cook 1997; Cuddy 2000; Goldsmith 2006; Pyburn 2003, 2008; Wille 2007; Wrobel 2003). This research design challenged the data access status quo by needing primary data for intersite comparisons, not the results of analysis that are found in journal publication. In addition, this research design calls for a broad range of expertise: geophysical surveyors, ceramicists, lithic experts, osteologists, hydrologists, epigraphers, etc. The list of specialists involved in a single investigation is ever increasing, as is the specificity of their language (Beebe 1998).

1996 Chau Hiix Artifact Inventory (Sample)

Data syntax: comma delimited, use x for missing data:

```

provenience,
material,
date (DD-MMM-YY),
count,
weight-in-ounces,
excavator-initials (slash/ separates up to 3 excavators),
#of# (number of bags),
notes

che-48-1-2-96,lithic,6-Mar-96,8,x,rls/kap,1,obsidian-flakes
che-48-1-20-96,ceramic,7-Mar-96,108,40,kap,1,treefall-cache
che-48-1-21-96,ceramic,7-Mar-96,103,44,x,1,treefall-cache
che-48-1-23-96,ceramic,7-Mar-96,26,32,gdw/rls,2of2,x
che-48-1-23-96,ceramic,7-Mar-96,68,38,rls/gdw,1of2,x
che-48-1-29-96,lithic,18-Mar-96,10,3,cjb,1,backdirt
bps-6002-1-1-96,ceramic,14-Mar-96,20,16,les,1,x
cht-97-1-4-95,carbon,20-Jul-95,x,x,lwb,1,x
bpt-93-1-5-96,ceramic,14-Mar-96,20,8,les,1,x
che-50-2-81-95,plaster,28-Jul-95,1,x,kap,1,painted
che-50-3-93-95,bone,30-Jul-95,1,x,gdw,1,stingray-spine
che-50-4-1-96,ceramic,22-Feb-96,1,x,jls,1,carved
che-70-4-110-96,lithic,26-Feb-96,1,17,hew,1,metate
chp-1-4-x-95,ceramic,19-Jul-95,1,x,pcc/dcc/hew,1,decorated
chp-112-1-12-96,bone,22-Jun-95,9,1,twc,1,faunal

```

FIGURE 2. 1996 Chau Hiix artifact inventory (sample).

To fulfill the research design, data collection and the management plan had to support data sharing across projects and a desire to make the primary data available for the next 100 years. Taking the very broad view that alphanumeric strings represent all data, building the digital Chau Hiix archive (CHARC) was approached as a problem of standardized representation. Being a small science project with basic digital technology expertise and limited finances, a bottom-up organization was developed based on the concept of Occam's razor: keep it simple. Thus, ASCII and TIF were selected as the basic archival formats.

The CHARC digital collection process began in 1996 as ADS was developing in England. The Archaeology Data Archive Project (ADAP) (Eiteljorg 1994), an ADS precursor in the United States, was publishing technology documentation that remains today a comprehensive introduction to digital archaeology (Eiteljorg 2008) and highly educational for a deeper understanding of applied technologies. At the same time, GIS software was migrating from arcane mainframe programs to desktop software, making it more accessible. The ability to import Chau Hiix primary data into GIS systems was identified as a goal for the analysis phase.

CHARC data are represented in spreadsheets, scanned versions of all primary notebooks, and thousands of well-pruned digital images, all documented in content-related README.txt files. Excavation activity at Chau Hiix ceased in 2007. Work on CHARC continues, however, to create the repository version: digitizing legacy data from 1990–1995, error checking, and the final pruning

of redundant files. Having monitored current repository guidelines, the data will then be ready for deposit and the 100-year access goal.

Digital Text, Spreadsheet, or Database?

Chau Hiix data collection involves collecting and maintaining data both in physical and digital form. Physical data consists of collected artifacts and field notes from excavators and surveyors. Digital data consists of spreadsheets for artifact inventory, survey points, personnel data, digitized field notebooks, and digital images. While the physical data is most valuable, issues of preservation and accessibility limit its potential shareability. The digital record provides for sharing Chau Hiix data both today and in the future through global digital repositories. Prior to 1996, Chau Hiix recordkeeping was on paper with pencil. Digital data collection began when a computer-typed record replaced pencil and paper, thus exchanging penmanship readability with typing errors. These first Chau Hiix digital records were keyed into a text editor, creating text (TXT/ASCII) files, the selected archival format. Data were entered in this text file with a documented comma-delimited syntax (Figure 2), mimicking the .CSV spreadsheet output format of today.

In 1996 it made sense to use a text editor to begin the Chau Hiix digital record due to software costs, lack of application expertise, and even lack of typing skills. By 2003 the technology skills of project personnel had standardized, software was more prevalent and affordable, and operating systems were more user friendly.

Data entry was done directly into a spreadsheet. The lab manager had to constantly monitor the staff with *greater* spreadsheet expertise, as they would make changes to the default spreadsheet settings to facilitate easier data entry, such as turning on auto-complete for cell values. The next person, with perhaps less experience, would not notice any auto-complete errors they were making. At the end of the day, the spreadsheet was checked for errors, exported into a comma-delimited ASCII file, both spreadsheet and ASCII files were copied to multiple external devices, and ASCII files were printed as an archival-backup version. That backup system continues today as CHARC is finalized for a repository.

During the years of digital data acquisition at Chau Hiix, there was an ongoing discussion about using a database instead of either a text editor or spreadsheet. Some small projects with technically savvy principal investigators, and many large well-funded projects with technology support staff, were developing database systems for managing their data. Technically savvy Chau Hiix students and staff offered to build databases for the project. But none of the core, permanent staff had the expertise to maintain, modify, or migrate a database for ongoing use over two decades of excavation. If a relational database system is not developed with excellent documentation and a regular maintenance plan, the data are in danger of becoming irretrievable. The risk of losing data, whether by lack of maintenance or data migration, is exemplified and well documented for the Newham Archive and others (Simpson 2004). Developing, maintaining, and troubleshooting a relational database requires technical expertise and regular attention. Chau Hiix chose a simple spreadsheet system with ASCII backups.

Prior to each field season, as project finances and expertise were evaluated, the same decision was made: spreadsheet data recording was more than adequate for field data acquisition, data-entry errors notwithstanding. Once documented and error checked, field data would never change. Therefore, ongoing change management as afforded by a relational database, was deemed unnecessary. It was also determined that the ASCII format was still the best archival format for data preservation. For a publishable archive of ASCII primary data, a relational database was not necessary.

A spreadsheet is basically a single table database with rows (records, entities, or elements) and columns (fields or attributes), often referred to as a flat-file database, like a phone book. All the data contained in a spreadsheet can be exported as a text file, with each row on a line and related column data separated by a *delimiter*, such as a comma, separating the column values, effectively the same as Figure 2. For field archaeology record-keeping, the spreadsheet is easy to create and tracks the simple lists of data being generated: what is found, where, when, and by whom.

While many projects now use relational databases in the field to manage their data collection, its advantages of managing changing data and multiple user access are not always essential in the field. Since archaeology records are never updated once free of data-entry errors, a small group of spreadsheets has continued to be an acceptable tool for data collection management. The spreadsheets form the foundation for potential data import into a relational database during analysis.

DESIGNING SPREADSHEETS AS DATA MODELING

The concept of data modeling describes the same process archaeologists engage in when planning for data collection in the field. For those with little database expertise, texts such as Sanders (1995) or Carlis and Maguire (2001) make data modeling accessible. It is like indexing and involves identifying entities (things with related properties), attributes (the properties), and the relationships between the two.

Data modeling can also be understood as an exercise in building a research design-specific ontology (Gruber 2007); that is, a catalogue of the things shared in a domain and particular context, and the relationships between those things. CHARC constructed four basic spreadsheets:

- Personnel (with cross-reference to field notebooks)
- Survey (points, activity proveniences, descriptions)
- Proveniences (excavation activity list, administrative information)
- Artifacts (general artifact inventory, small finds identification)

For each spreadsheet (Figure 3), column headings are spelled out and include the syntax when appropriate. When journal articles publish relational database schemas without documenting the attribute codes, the relationships between different tables are often too cryptic for the reader to understand or use as a model (Figure 4).

Data Redundancy as a Spreadsheet Strategy

Before addressing the details of specific spreadsheet attributes, note the data redundancies in Figure 3. Row 1 indicates redundant “year” information across the spreadsheets. In row 2, the “initials” field of the personnel spreadsheet serves as a look-up for the initials recorded in the other three spreadsheets. In row 3, each provenience record is placed in the geophysical survey context through its unique point number. In rows 4–8 every artifact record is connected to a provenience record, which in turn is connected to a survey point. Notebook-number in row 7 links a specific provenience to the field notes of the assigned excavator. Note the **bold** text in rows 8–9, which indicates the need for controlled vocabulary.

Data redundancy is a strategic principle for Chau Hiix data collection. There is no data redundancy in a relational database, so when a piece of data changes, such as a customer moving to a new address, it is changed in only one place. Since the primary data in archaeology will not change over time, it turns out that during data collection spreadsheet redundancy is a positive situation. At the analog data collection stage, where artifacts are being dug from the ground, placed in bags, processed, and conserved in the lab, there are multiple opportunities for recording error before the digital record is created. These early, non-digital recording errors do not become evident until sorting the digital record.

The spreadsheet sorting tools readily identify data consistency problems. Consider the redundancy of “year” in each

	Personnel	Proveniences	Survey	Artifacts
1	year (YYYY)	year (YYYY)	year (YYYY)	date collected (YYYY-MM-DD)
2	initials (3)	lead excavator initials	surveyor initials	excavator initials
3	name-last	datum point number	point number (unique)	small find number (if applicable)
4	name-first	location (CH/BP) type (ETPS)	northing	provenience type
5	name-middle	location number (unique)	easting	provenience number
6	home (country/state/school)	location operation number	elevation	provenience operation number
7	notebook number (n-YY##)	notebook number(s)		provenience context number
8	job (use standard terms)	structure number	descriptive code	provenience year (YY)
9	status (use standard terms)	date begun		material type (use standard terms)
10		size in meters		description from field (short text)
11		cross-ref proveniences		count
12				weight (ounces)
13				number of bags (if applicable)
14				repository location (use codes)
15				accession number (unique)
16	notes	notes	notes	notes

FIGURE 3. Chau Hiix data collection basic spreadsheets with colors indicating data redundancies.

spreadsheet. Chau Hiix had a backlog of artifact accessioning due to the volume of material recovered each year. Each season, data recording staff would have to pay close attention to the specific year on the bag/lot being accessioned. Errors easily happened. Data-entry errors were confounded by excavator error when writing the provenience number on the artifact bag or overfilling the bags. In the lab, overfilled bags were re-bagged in manageable sizes, creating another opportunity for pre-digital data-entry error.

Excavator initials represented those making the actual data discovery choices. The lead excavator often had their crew take responsibility for collected material; thus, there would be multiple initials on each bag. Crewmembers were inconsistent in the recording order of initials and redundancy across spreadsheets allowed for clarification when researching field notes for discrepancies.

The redundancy in the data representation and entry process has made it possible to do error checking when there is an information discrepancy between the physical artifact, the notebook references, photographs with incomplete provenience references, lost surveyor locations, and even which staff were on site any given year. Redundancy across spreadsheets thus supports a chain of evidence model; that is, a digital archive containing

the record of all activity, from the moment the earth is disturbed, through all stages of processing and analysis, until the resulting information/data is deposited in the final, publishable archive. At that point, data redundancy can be removed only if it served the archive-building and error-checking process.

Survey and Provenience Spreadsheets

Taken together, the survey and provenience spreadsheets are at the heart of data collection for a given research design. The research design identifies what to look for, the survey identifies the real-world context, and the provenience designation documents the specific activity locations. The surveyor performs the topographic survey, identifies geophysical features, and assigns each feature a unique number. When the project director assigns an activity number (provenience), the surveyor ensures that there is also a geophysical designation with a datum point. For the Chau Hiix project, all activity, material, and documentation are tied to the survey through the provenience designation. Every provenience number refers to a specific location. Every object removed from an excavation activity, either on or within the earth, is assigned a provenience number. In the case of "small finds" (objects of special interest such as complete vessels, jade, shell carvings, etc.), the provenience number is written directly on the object.

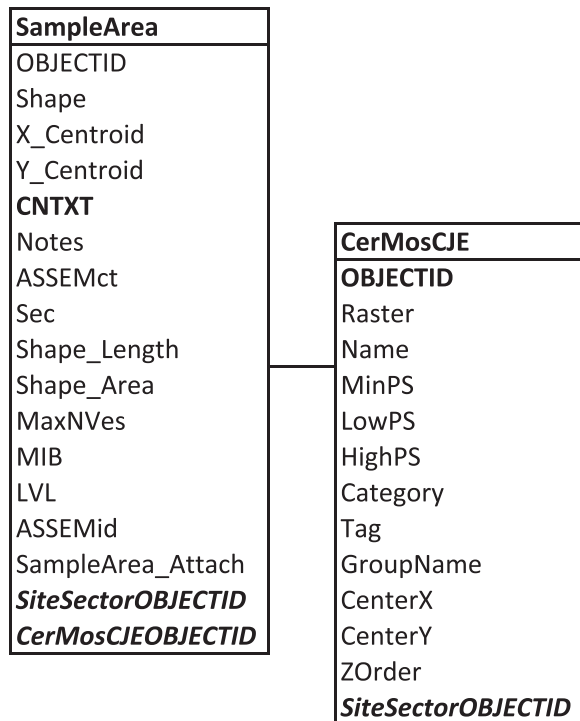


FIGURE 4. Example of a relational database schema derived from Sharp and Litschi (2014:Figure 3).

Provenience Numbers

The survey data begins a chain of evidence that links all other data through a well-constructed provenience numbering system. The provenience number is made up of six parts: location, activity type, activity number, operation, context, year (see examples in Figure 2). The general site location accommodates the issue of outlier areas of interest that are not contiguous with the area of focused activity. CH represents the main site, but to determine activity relationships there were several surveys, surface collections, and test pits in more remote areas, outliers such as Ben's Pass (BP).

There are four activity types: surface collection, posthole, test pit, and excavation. The surface collection (S) is either an area or a point provenience for any surface collected artifacts. There were several investigations that involved postholes (P), and this also accommodated any postholes created to construct buildings or living quarters. Test pits (T) were defined as preliminary investigations no larger than 1 × 1 m. Anything larger was considered a full Excavation (E). Each activity (E, T, P, or S) was given a unique number assigned by the lab manager. Activity numbers are never repeated and are tied to a specific datum as provided by the surveyor.

The operation number for each activity is a sequential number beginning with one (1). If an activity in a specific location spans more than one year, the second year becomes operation two (2), as when an excavation is reopened in another field season. Within each field season there is always the potential that a specific activity needs to be expanded from its initial size des-

ignation. In that case the expansion becomes a new operation number. For example, in 1994, excavation CHE-50 operation 1 (CHE-50-1) began in the main pyramid. In 1995, CHE-50 was reopened and designated CHE-50-2. A tomb was discovered but the crypt extended beyond the confines of the excavation shaft. Therefore, a new operation was opened, CHE-50-3, and a contiguous shaft was excavated until the crypt was reached. So, an operation number represents activity in a specific location, or contiguous locations, but is not necessarily tied to year. The redundancy of the year component helps clarify recordkeeping errors.

The context number is assigned by the excavator and can be used to identify material or features. For example, in CHE-50-2 the initial observation of the crypt was given its own context number, CHE-50-2-45. The meaning of the context number is documented in the excavator's notebook. The excavator may ask the surveyor to point provenience any given context. The surveyor shoots a point with a unique point number and uses the provenience number as the description.

Finally the provenience number ends with a two-digit year designation. Again, the redundancy helps with error checking and maintaining the chain of evidence.

The Problem with Features

Informed by the research design hypothesis, the archaeologist selects a specific x-y-z location in the world to excavate. As the digging begins, the excavators record their observations. These observations, however, are biased interpretations based on knowledge and experience. The more complex the area of excavation, the more interpretation can come into the observation record. An uncovered layer of broken pottery could be interpreted as "features," such as the top of a midden, an indication of a hearth, burial goods, or simply a random broken pot that in later analysis indicates the presence of manufacturing. All are viable potential interpretations that must be considered in order to inform the ongoing observation and excavation technique. At what point does the observation (primary data) become interpretation (secondary data)? Each investigator must decide that for themselves, yet the recordkeeping system should be as focused on the primary data as possible.

The key to the Chau Hiix provenience representation system is that all observations and material are given individual provenience numbers, since the concept of "feature" is considered a construct of the excavator and represents an interpretation that may or not prove valid in the future. With individual proveniences, collected materials can be grouped and regrouped as more observations occur over multiple operations. For example, after the burial recovery in CHE-50-2 and CHE-50-3, it was discovered that there was an even earlier burial underneath. Initial field assumptions assigning grave goods to specific burials were corrected during the analysis phase. Burials are recognized as a "feature," but all associated materials are catalogued as independent objects. Since every artifact has its own number, it is easy to construct and deconstruct features or assemblages (Figure 5). In the CHE-50-2 crypt example, the 12 blades beneath the skeleton were each given individual context numbers (Figure 6). Features and assemblages do not fit the concept



FIGURE 5. Post-collection assemblage of stingray spines with small find numbers.

of Occam’s razor and thus are minimized in the primary data by assigning individual object provenience numbers.

Not every artifact receives its own number. Excavators must still make field determinations about what data should be treated as a single piece of information. For example, artifacts in a midden that were deposited as a single event may be given a single provenience. Diagnostic artifacts from such contexts may be indicated with the addition of a small find number that becomes an additional data point in the artifact spreadsheet.

Ultimately, the purpose of this simplified notational system is to make the data as amenable as possible to reanalysis by future researchers. The aim is to reduce the number of assumptions made about an assemblage at the time of recovery that cannot be altered later, and to ensure that the assumptions that were made are explicit so that limits to possible reanalysis of the assemblage are clear.

Artifact Spreadsheet

The artifact spreadsheet is the inventory of collected materials. In the field, the excavator puts collected material into bags separated by material type. Bags are labeled with provenience, excavator initials, and date and sent to the lab. A given context may have just a single ceramic piece, or there may be so much ceramic material that it fills multiple bags. In the lab, each bag is considered to be a data record in the spreadsheet. Every artifact data record is assigned an arbitrary, unique accession number. This number supports analysis phase logistics when there are multiple bags from a single context: some contexts generated hundreds of individual ceramic pieces. Bag numbers help relocate specific pieces in the physical collection. Accession numbers facilitate analysis by multiple analysts, supporting bag for bag interpretation comparisons.

Note the redundancy in the fields “date-collected” and “provenience-year.” While the date-collected field could serve as the data access point for “year,” the inclusion of the “provenience-year” field allows for the collocation of artifacts from different collection dates and contexts in order to create assemblages.

FILENAMES

Before determining the values that populate individual spreadsheet cells, consider how the provenience number can be used to name and collocate the many digital files that are created during data collection. Names are needed for directories, folders, documents, and artifact images. The computer processes these names as an alphanumeric string and is not concerned with whether filenames are meaningful or non-meaningful. Non-meaningful filenames are unique random numbers that reference each file, such as the names generated by digital cameras. These are not user-friendly when contemplating a directory structure and require a file management system that is easily managed with a relational database but difficult to establish with a spreadsheet.

Meaningful (descriptive) filenames allow the researcher access to data by browsing. Repositories are clear about the structural format of filenames: characters A–Z and 0–9, no spaces, etc. However, there are no examples and little guidance for how meaningful names are actually constructed (e.g., Burke et al. 2009 or Lloyd-Smith 2016). Meaningful filenames might include date, location, excavator initials, or any of the fields in the CHARC spreadsheets. CHARC uses the provenience number as the starting point for file naming (Beebe 2010) so that data can be assembled and reassembled based on this most basic attribute that links all other data and begins the chain of evidence.

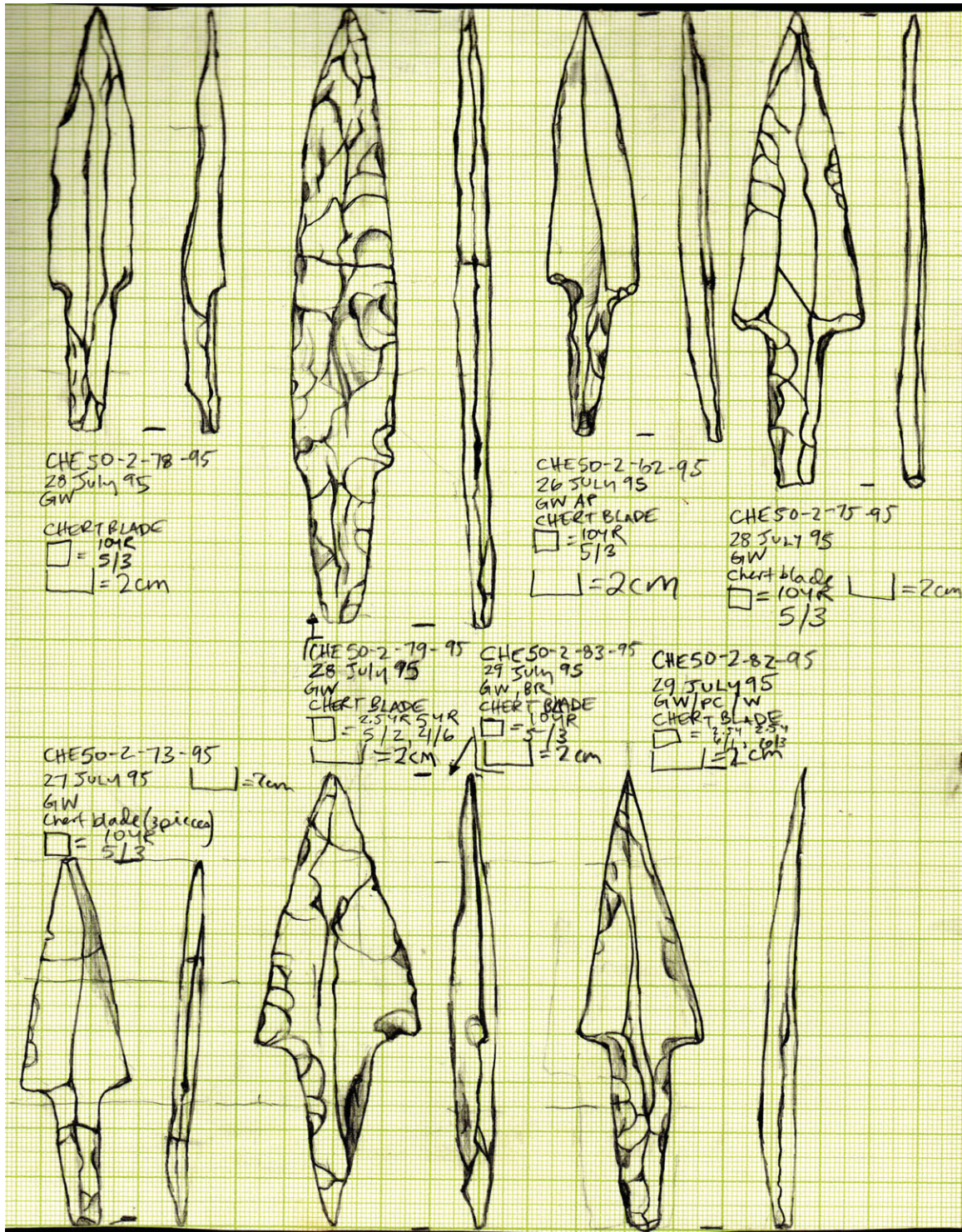


FIGURE 6. Tomb 1 lithics with individual context numbers.

The filenames for Chau Hiix excavation and survey forms, images (photos, drawings), burial data, etc., all include the provenience number, either partial or complete, depending on the level of reference required (Figure 7). Burial, structure, and small find numbers serve as provenience shorthand for a specific object

or location. Within each directory folder, therefore, the files are sorted by provenience. Thus, during analysis, the spreadsheets can be sorted by the attribute of interest, and then the associated provenience value is used to locate specific files.

File/folder names	Description
Folder: 9402-IAW	1994 field notebook #2, excavator initials (IAW)
9402008.pdf	Pages 1 through 8
9402016.pdf	Pages 9 through 16
Folder: Excavation forms	F indicates Form, CH implied unless otherwise noted
F-BPE-1-98.pdf	Ben's Pasture (BP) excavation 1, 1998
F-P112-96.pdf	Posthole 112, 1996
F-T159-99.pdf	Testpit 159, 1999
F-S2039-97.pdf	Surface collection 2039, 1997
Folder: Burial data	Burial number precedes provenience
B11-E6-93.pdf	Burial 11 in excavation 6, 1993
B45-E50-5-30-96-plan.tif	Burial 45 in excavation 50, 1996
Folder: Images	Codes: P-photo, M-Structure Number, D-drawing
P-E18-3-11-03-p3.tif	CHE-18-3-11-03, photo3
P-M7-E22-5-13-05-p11.jpg	M7, CHE-22-5-13-05 (2005), photo 11 of plan views
D-sf10618.tif	Drawing of smallfind #10618
P-sf10618.tif	Photo of smallfind #10618
P-sf10618.jpg	Photo of smallfind #10618, in situ (.jpg)
Folder: Survey Forms	F indicates Form, M is structure number, year (YY)
F-M182-97.pdf	M182 surveyed in 1997
F-M182-03.pdf	M182 re-surveyed in 2003
Folder: Area maps	A indicates surveyor drawings, sketches, point printouts
A-BP-96-006.tif	Area map of Ben's Pasture, 1996, sketch 6
A-M1-98.tif	Area map of M1, 1998 (CH implied)
A-M13-05.tif	Area map of M13 (platform of M1, other M's), 2005

FIGURE 7. Implementation of CHARC filename system.

CONTROLLED VOCABULARIES

Consistency in the alphanumeric strings that constitute the values in each cell of a spreadsheet is at the heart of data representation in this simple, spreadsheet-based data collection plan. These consistent alphanumeric strings take the form of controlled vocabularies, such as provided in the Art & Architecture Thesaurus (AAT) (2015), or the prescribed type of numeric values. In the broadest sense, the allowable form of a numeric value can be considered a type of controlled vocabulary. Having control over these alphanumeric strings is essential when working with spreadsheets. It facilitates the search and discovery process, as well as the organization of files with meaningful names.

Personnel Spreadsheet

Consider how the personnel spreadsheet (Figure 3) exemplifies a basic tenet: every cell contains only one piece of data. In spreadsheet practice, the name of each person is represented in three separate fields: first, middle, and last. It is commonly understood that entering first-middle-last name in a single cell will result in a sorted list of all the John or Mary names grouped together, which is less useful than a Smith (John) and Smith (Mary) grouping. Note that in the "home" field, the three pieces of data are grouped together with a slash (/) delimiter and a required ordering. This exemplifies a controlled shortcut for entering data in the field. The personnel "home" data is not essential for fieldwork, but using spreadsheet tools the delimiter allows for expansion into three columns for the spreadsheet's final archival version.

The "status" and "job" fields exemplify the need for controlled vocabulary. Since the Chau Hiix project has spanned more than 25 years, keeping track of personnel and their status at the time of project participation became an issue. The evaluation of notebook data and observations might be impacted by knowing that undergraduate student A in 1996 participated as a graduate student in 1998, as a PhD candidate in 2000, and a professional archaeologist in 2005. Similarly, after 25 years and hundreds of participants, being able to sort by various jobs was helpful when considering specialists to contact.

Facet Analysis

The personnel job field exemplifies the result of facet analysis to generate the controlled vocabulary for the job attribute (Figure 8). A facet analysis is a bottom-up categorization method (Hedden 2016; Ranganathan 1967) that considers what exists in a collection (or ontology), analyzes the component parts (characteristics), groups the components into related categories, then establishes relationships both between and within the groups. This is in contrast to an enumeration method that attempts to list all possible groups that exist, have ever existed, or may exist. A faceted scheme is initially not hierarchical, so it easily allows for the addition of new attributes or values. A well-constructed faceted scheme builds as need arises. Enumeration, as a top-down method, requires that all categories preexist, thus when a new category emerges there are problems putting it into the scheme. For example, the Library of Congress Classification (LCC) scheme is enumerative in that it provides for every book

Job analyst (type: list 1) archive processor artist excavator (type: list 2) foreman manager (type: list 3) map (type: list 2) photography specialist (type: list 4) staff (hourly) visitor	Status FS (field school-undergraduate) MA (Masters student) PHD (Graduate student) Professional
List 1: Analyst types osteology architecture ceramics lithics iconography hydrology fish bone (add as needed)	List 2: Location types main site outlier settlement List 3: Manager types archive lab project List 4: Specialist types GIS remote sensing (add as needed)

FIGURE 8. Controlled vocabulary for personnel job and status.

to have a call number that places that book on one and only one shelf location. When “computer science” emerged as a discipline in the mid-twentieth century, it was forced into the LCC class “QA Calculating Machines,” which was a subclass of mathematics (Miksa 1984:65).

Rather than anticipating and enumerating every possible personnel job, the existing jobs were grouped and their characteristics identified. Characteristics were then grouped into categories that could be applied to multiple jobs. For example, rather than enumerating every possible excavator job (e.g., main site excavator, settlement excavator), the main concept is excavator and is qualified by location terms (Figure 8, List 2). This not only allows for a person to be a general excavator in multiple locations, but it also provides for the addition of other locations as qualifiers, should that situation arise. This is basically the same approach used in data modeling (Sanders 1995) and was applied later when developing material types (Table 1).

Both spreadsheets and relational databases allow for facets to be combined in various orders depending on the researcher’s focus. In the CHARC artifact spreadsheet, the ceramicist sorts by material type first and then considers all the other fields (columns). Someone studying burials might sort first on small find number in order to get a subset of grave goods. The enumerative

approach is to lump artifacts together under the feature concept. The faceted approach splits feature objects into individual representations for collocating later based on expert analysis. The artifact spreadsheet collects feature interpretations in the “field description,” but multiple interpretations are facilitated by the individual representation of every object.

Two other values used facet analysis to determine controlled terms: survey codes and material types. Problems establishing survey codes and syntax are documented elsewhere (Beebe 1998) but resulted in a simple list of elements (topo point, station point, mound number, provenience number) with sub-attributes of either “datum” or a directional reference (NE, SE, SW, NW, top, center, bottom). Descriptors were added to indicate feature interpretations (e.g., aguada). The problem of multiple parts to a “building” was addressed in the same way as any feature. Each part was given its own M-mound number, so as investigations progressed complex structures could be created out of the component parts, assigning a new M#. Flood et al. (1989) provide more details for exploring survey description methods.

Prior to developing material type terms, the example of 12 artifacts in the facet analysis exercise (Figure 9) exemplifies the “field description” in the artifact spreadsheet. The field descriptions proved impossible for sorting, as did uncontrolled material

Attribute elicitation

1. List at least 12 potential artifacts for field collection
2. Group artifacts into threes and identify common attributes
3. Swap out one object for another from the twelve
 - a. Identify common attributes again
 - b. List attributes that exist, may exist, or have existed
 - c. Repeat replacing one object and identify common attributes. Triad comparisons help identify less obvious attributes.
4. From the attribute list, identify logical opposites, for example: Decorated and not-decorated
5. Create logical groups of attributes. Groups must be mutually exclusive: no two-word attributes (note 'groundstone' becomes a one word term)
6. Select an authorized term to use when attributes are synonyms.

Example of 12 potential artifacts for Chau Hiix collecting

1. carved jade beads
2. broken chert blades
3. obsidian debitage
4. groundstone metate fragment
5. hematite earspool
6. netweights or mendhole sherds
7. carved bone flute
8. worked conch shell
9. animal skeleton
10. bones in an ancestor basket
11. pomacea
12. polychrome vessel

FIGURE 9. Facet analysis for material types.

names. For example, when the lithic analyst sorted the inventory, she had to look at groupings for G-groundstone, H-hematite, O-obsidian, C-chert, and J-jade. The controlled vocabulary (Table 1) that evolved from the materials facet exercise resulted in the use of the general term lithic, with groundstone, jade, and chert as a subset of mutually exclusive terms.

Data attributes fall into three basic categories (Table 2): physical, conceptual, and contextual (Beebe 2006:4). Physical attributes are the focus of data repositories and include both administrative (e.g., access rights, repository, copyright, file type) and biographical characteristics (e.g., excavator, site). Conceptual are both naming (e.g., lithic, ceramic) and interpretation (e.g., ear-spool, metate). Contextual provides for the relationships among objects (e.g., cooking, debitage type).

Conceptual and contextual attributes use controlled vocabularies as attribute values. CHARC relied on material type analysts to develop interpretation and contextual attributes, though excavator interpretations were captured in the field description. These were often corrected or clarified by the material specialist in the analysis phases. Analysis vocabularies are for item-level description and access. The AAT exemplifies a controlled vocabulary that could initially be used for archaeology, the lack of archaeological detail notwithstanding.

The AAT is a facet-like thesaurus that holds potential for item-level description but lacks the breadth and depth required for archaeological items. For example, the AAT has no terms for describing the details of a biface such as bi-pointed, macro-blade, or eccentric. Analysis level vocabulary is a starting place

in archaeology for material specialists to build standardized vocabulary. *The Archaeologist's Fieldwork Companion* (Kipfer 2007) provides preliminary vocabulary lists, some faceted, some enumerative, and some needing to be processed into facets or hierarchies.

In practice, most classification systems are a combination of enumerative and faceted categorization. Facet analysis creates subtypes that provide for various combinations of attributes. The facet structure results in attributes and values that can be ordered (enumerated) for a particular focus of research; for example, the sorting order used in the spreadsheet.

Attributes as Metadata

Metadata is a structured representation of information types that facilitates retrieval. Metadata uses ontologies to describe things. Dublin Core (DC) (2015) and VRA Core (Visual Resource Association 2007) are ontologies with metadata schemes composed of the three basic attributes. DC was designed as the minimum set of attributes for representing an object in the digital environment (Clarke 2015:Table 1). While DC is for general resources, VRA Core is specialized for images and cultural resources. Both emerged in the mid 1990s, VRA forming core elements that crosswalked to DC (Table 2). VRA recognized special needs with regard to locating and describing visual resources. For example, the three panels of Monet's triptych *Agapanthus* are located in three different art museums: St. Louis, Kansas City, and Cleveland. DC did not accommodate such a collocation problem. This is similar to archaeology's issues with features and assemblages. Consider, also, VRA's expansion of the DC Extents element to

TABLE 1. Vocabulary from Facet Analysis of Material Type.

Material Types	Sub-type	Sub-sub-type
bone	faunal	stingray spine turtle (etc.)
bone	human	
carbon		
ceramic	disc mendhole netweight	
ceramic	vessel	partial whole
historic	ceramic glass metal	
lithic	cinnebar chert ground stone hematite jade obsidian	
misc	coral pumice wood	
plaster	daub plastered stone stucco	painting
shell	fresh	clam jute mother-of-pearl mussel pomacea
shell	marine	conch olivella spondylus
soil	chemical flotation waterscreen	

VRA's Measurement element with attributes more specific to image needs (Table 3).

For archaeology, the elements required for the tDAR repository serve as a specific instantiation of Dublin Core. tDAR documentation on creating and editing resources provides for the creation or import of ontologies as well as adoption or crosswalking to ontologies already deposited by other projects. These crosswalk opportunities facilitate the research community building ontologies with shared attributes and values from the bottom up, in the fashion of facet analysis. A focused effort by the domain of archaeology at this level of terminological standardization would facilitate item-level discovery and eliminate the need for each individual researcher to create vocabularies or crosswalks.

TABLE 2. Attribute Types and Metadata Scheme Crosswalk.

Attribute Type	Dublin Core	VRA Core
Administrative	Type	Work, collection or image
Biographical	Creator Contributor	Agent
Contextual		CulturalContext
Administrative	Date Created	Date
Contextual	Description	Description
Contextual	Description	Inscription
Administrative	Subject or Coverage Spatial Resource Identifier Source	Location
Biographical		Material
Administrative	Format Extent	Measurements
Contextual	Relation Source	Relation
Administrative	Rights	Rights
Biographical		Source
Biographical		StateEdition
Contextual		StylePeriod
Conceptual	Subject or Coverage Temporal	Subject
Biographical	Format Medium	Technique
Administrative		Textref
Biographical	Title	Title
Biographical	Type	worktype

As individual archaeologists embrace the need to curate their digital data and the structural requirements for data reuse, consideration should be given at the data collection planning stage for the ontologies, attributes, and controlled vocabularies that will be needed. Kulasekaran et al. (2014:61) explored a metadata model for archaeology to illustrate "that DC metadata can act as an integrative platform for a non-traditional (but increasingly common) researcher-curated, distributed repository environment." CIDOC CRM Core emerged in 2006 as a standard for semantic interoperability in a broadly defined cultural heritage domain (International Committee for Documentation 2006). CRM-Core (Conceptual Reference Model) is an ontology with an extensive set of entities, is based on the attribute of "event," and can be crosswalked to DC. VRA Core and CRM Core are both at the foundation of ArchaeoCore (2015) emerging in the art domain (Brenz et al. 2013). Archaeologists might consider adapting and adopting this effort.

THE CHARC EXAMPLE

The CHARC data collection plan (Figure 1) demonstrates the functionality of simple spreadsheets. Spreadsheets are easily maintained, require minimal technology skill, and are affordable for small science projects. When relational database training

TABLE 3. Definitions for the Measurement Element from the VRA Core.

Type	Definition
area	
base	Base value used to determine scale along with a target value.
bitDepth	In a digital file, the number of colors for an image; calculated as 2 to the power of the bit depth; for example, a bit depth of 8 supports up to 256 colors, and a bit depth of 24 supports up to 16 million colors.
circumference	The length of the boundary line of a circle, figure, area, or object.
count	Number of units.
depth	The extent, measurement, or dimension downward, backward, or inward.
diameter	The length of a straight-line segment passing through the center of a figure, especially of a circle or sphere, and terminating at the periphery.
distanceBetween	The length of the distance between two points.
duration	The length of a period of existence or persistence, as in a performance.
fileSize	The length of a digital file in bytes.
height	The distance from the base of something to the top.
length	The measurement of the extent of something along its greatest dimension.
resolution	The degree of sharpness of a digital display, expressed as a matrix of dots. For example 1024 × 768.
runningTime	Total length of time it takes to play back recorded time-based works such as film or video.
scale	The ratio between the size of something and a representation of it. For example ¼"=1'.
size	Any of a series of graduated categories of dimension whereby manufactured articles, such as shoes and clothing, are classified.
target	Target value used to determine scale along with base value.
weight	A measure of the heaviness of an object.
width	The measurement of the extent of something from side to side.
other	When none of the other type attributes apply, use other and where possible specify the other value in the notes sub-element.

and technology become as easy to use and ubiquitous as word-processing, then the data management plan can begin with a relational database tool. Any database tool, whether flat or relational, begins with the same problems of data communication outlined here in the data collection plan.

Communicating for future digital data sharing is a problem of vocabulary and organization. While resources and repositories call for organizational structure and naming conventions, the Chau Hiix vocabularies and provenience number provide an example of data representations. There is no "right" way to organize or represent data. A representation must reflect the scope of the content and the objectives of the user. The phone book is alphabetical, the TV guide by time, the QWERTY keyboard by frequency and speed.

The CHARC table of contents (Figure 10) is the result of a facet analysis and for 25 years has functioned well as the organizing scheme for Chau Hiix data collection and represents all data that will eventually reside in a repository. The faceted approach to organization supports the ability for researchers to access the data from their particular research interest. Analysis level vocabulary (conceptual and contextual) is a starting place in archaeology for material specialists to begin building standardized vocabulary. AAT provides a foundation, but needs archaeological detail, like introducing "provenience," which currently in AAT defaults to the broader term "provenance."

CHARC : Chau Hiix Archive
Table of Contents

- Readme Introduction
- 1. Manuals
- 2. Personnel
- 3. Notebooks
- 4. Proveniences
- 5. Artifacts
- 6. Survey
- 7. Visuals
- 8. Analyses
- 9. Documents
- 10. Seasonal files archive

FIGURE 10. Chau Hiix Archive (CHARC) table of contents.

Creating standardized terminology to populate the values in the spreadsheets is a time-consuming process. Future archaeologists would benefit from a community effort to develop a thesaurus of archaeology terms using AAT as a starting place. Establishing common vocabularies and systems for numbering and collecting among the various site-based explorations would facilitate the promise of archaeological science in its ability to reproduce

and support each other's theories and potentially create some knowledge about the past.

While ADS pioneered the field of data repositories, today there are many other repository projects to consider, Open Context (2007) being one that fits well with the CHARC spreadsheet approach. A focused effort by the archaeology community to create a clearinghouse of repositories, and to consider adoption of a specific metadata standard and controlled vocabulary governing attribute values, would support the usability, interoperability, and discovery of currently disparate data sets.

Data Availability Statement

No new data were used in this paper. When completed, CHARC will be available from the author or Dr. Anne Pyburn, Department of Anthropology, Indiana University, Bloomington (apyburn@indiana.edu).

REFERENCES CITED

- Andres, Christopher R.
2005 *Building Negotiation: Architecture and Sociopolitical Transformation at Chau Hiix, Lamanai, and Altun Ha, Belize*. PhD dissertation, Department of Anthropology, Indiana University, Bloomington.
- Andres, Christopher R., and K. Anne Pyburn
2004 *Out of Sight: The Postclassic and Early Colonial Periods at Chau Hiix, Belize*. In *The Terminal Classic in the Maya Lowlands: Collapse, Transition, and Transformation*, edited by Arthur A. Demarest, Prudence M. Rice, and Don S. Rice, pp. 402–423. University of Colorado Press, Boulder.
- ArchaeoCore
2015 Institute of Fine Arts, NYU. Electronic document, <http://www.ifaresearch.org/archaeocore/>, accessed April 20, 2017.
- Archaeological Data Services/Digital Antiquity (ADS)
2011 Guides to Good Practice. Electronic document, <http://guides.archaeologydataservice.ac.uk/>, accessed November 6, 2016.
- Art & Architecture Thesaurus (AAT)
2015 The Getty Research Institute. Electronic document, <http://www.getty.edu/research/tools/vocabularies/aat/>, accessed November 8, 2016.
- Beebe, Caroline
1998 *Joining the Club: Issues, Problems and Practices Initiating GIS*. In *New Techniques for Old Times: CAA 98—Computer Applications and Quantitative Methods in Archaeology*, edited by Juan A. Barcelo, Ivan Briz, and Assumpció Vila, pp. 89–95. BAR Archaeology Series 757. Archaeopress, Oxford.
- 2006 *Bridging the Semantic Gap: Exploring Descriptive Vocabulary for Image Structure*. PhD dissertation, School of Library and Information Science, Indiana University, Bloomington. Electronic document, https://ils.unc.edu/~wildem/ASIST2007/Beebe_dissertation.pdf, accessed March 3, 2017.
- 2010 *The Chau Hiix Archive: Principles, Problems, and Solutions*. Paper presented at the 75th Annual Meeting of the Society for American Archaeology, St. Louis. Electronic document, <http://alexandriarchive.org/blog/wp-content/uploads/2010/04/BEEBE.ChauHiixDigitalArchive.pdf>, accessed November 9, 2016.
- Berez, Elizabeth, Ann R. F. Burns, and Lucie Wall Stylianopoulos
2013 *Cataloging in the Cloud: Shared Shelf and ArchaeoCore*. *VRA Bulletin* 39(2/2). Electronic document, <http://online.vraweb.org/vrab/vol39/iss2/2>, accessed April 20, 2017.
- Burke, Heather, Claire Smith, and Larry J. Zimmerman
2009 *The Archaeologist's Field Handbook: North American Edition*. Alta Mira Press, Lanham, Maryland.
- Carlis, John, and Joseph Maguire
2001 *Mastering Data Modeling: A User-Driven Approach*. Addison-Wesley, Boston, Massachusetts.
- Clarke, Mary
2015 *The Digital Dilemma: Preservation and the Digital Archaeological Record*. *Advances in Archaeological Practice* 3:313–330.
- Cook, Patricia M.
1997 *Basal Platform Mounds at Chau Hiix, Belize: Evidence for Ancient Maya Social Structure and Cottage Industry Manufacturing*. PhD dissertation, Department of Anthropology, University of Arizona, Tucson.
- Cuddy, Thomas W.
2000 *Socioeconomic Integration of the Classic Maya State: Political and Domestic Economies in a Residential Neighborhood*. PhD dissertation, Department of Anthropology, Columbia University, New York.
- Digital Antiquity
2013 *The Digital Archaeological Record (tDAR)*. Electronic Document, <https://core.tdar.org>, accessed November 9, 2016.
- Dublin Core Metadata Initiative
2015 *Metadata Basics*. Electronic document, <http://dublincore.org/metadata-basics/>, accessed November 9, 2016.
- Eiteljorg, Harrison
1994 *The Archaeological Data Archive Project (ADAP)*. Electronic document, <http://csanet.org/archive/adap/adaplond.html>, accessed November 5, 2016.
- 2008 *Archaeological Computing*. 2nd ed. Electronic document, <http://archcomp.csanet.org>, accessed November 5, 2016.
- Flood, Josephine, Ian Johnson, and Sharon Sullivan (editors)
1989 *Sites and Bytes: Recording Aboriginal Places in Australia*. Australian Government Publishing Service, Canberra.
- Goldsmith, A. Sean
2006 *The Houselots of Chau Hiix: A Spatial Approach to the Study of Non-Elite Domestic Variability at a Small Maya City*. PhD dissertation, Department of Archaeology, University of Calgary, Alberta.
- Gruber, Tom
2007 *Ontology*. Electronic document. <http://tomgruber.org/writing/ontology-definition-2007.htm>, accessed June 21, 2017.
- Hedden, Heather
2016 *The Accidental Taxonomist*. 2nd ed. Information Today Inc., Medford, New Jersey.
- International Committee for Documentation
2006 *CIDOC CRM*. Electronic document, <http://www.cidoc-crm.org>, accessed March 1, 2017.
- Kipfer, Barbara Ann
2007 *The Archaeologist's Fieldwork Companion*. Blackwell Publishing, Malden, Massachusetts.
- Kulasekaran, Sivakumar, Jessica Trelogan, Maria Esteve, and Michael Johnson
2014 *Metadata Integration for an Archaeology Collection Architecture. Proceedings of the International Conference on Dublin Core and Metadata Applications. Austin*. Electronic document, <http://dcpapers.dublincore.org/pubs/article/view/3702/1925>, accessed March 3, 2017.
- Lloyd-Smith, Lindsay
2016 *Archaeology: Principles and Practices of Digital Data Management*. In *Bringing Digital Data Management Training into Methods Courses for Anthropology*, edited by Blenda Femenias. American Anthropological Association, Arlington, Virginia. Electronic document, <http://www.americananthro.org/LearnAndTeach/Landing.aspx?ItemNumber=20641&navItemNumber=20708>, accessed November 8, 2016.
- Miksa, Francis
1984 *The Development of Classification at the Library of Congress. University of Illinois Occasional Paper No. 164*. Electronic document, <https://www.ideals.illinois.edu/bitstream/handle/2142/3957/gslisoccasionalpv00000100164.pdf>, accessed November 8, 2016.
- Open Context
2007 *The Alexandria Archive Institute*. Electronic document, <https://alexandriarchive.org/what-we-do/open-context/>, accessed March 1, 2017.

- Pyburn, K. Anne
2003 The Hydrology of Chau Hiix. *Ancient Mesoamerica* 14:123–129.
2008 Pomp and Circumstance before Belize: Ancient Maya Commerce and the New River Conurbation. In *The Ancient City: New Perspectives on Urbanism in the Old and New World*, edited by Joyce Marcus and Jeremy Sabloff, pp. 247–272. National Academy of Sciences, Washington, DC; School of American Research, Santa Fe, New Mexico.
- Ranganathan, Shiyali Ramamrita
1967 *Prolegomena to Library Classification*. 3rd ed. Ranganathan Series in Library Science Vol. 20. Asia Publishing House, London.
- Sanders, G. Lawrence
1995 *Data Modeling*. Boyd & Fraser, Danvers, Massachusetts.
- Sharp, Kayeleigh, and Melissa Litschi
2014 Maximizing E-Data Collection: A Novel Approach for Data Collection and Transmission Using Tablet Technology. *Advances in Archaeological Practice* 2:104–122.
- Simpson, Duncan
2004 *Risk of Loss of Digital Data and the Reasons It Occurs: Case Studies and Analysis*. Digital Preservation Coalition. Electronic document, <http://www.dpconline.org/docman/miscellaneous/members/117-risk-of-loss-of-digital-data/file>, accessed March 1, 2017.
- Visual Resource Association (VRA)
2007 VRA Core 4.0. Electronic document, <http://www.loc.gov/standards/vracore/>, accessed November 9, 2016.
- Wille, Sarah J.
2007 Sociopolitics and Community-building: The Entanglement of Pre-Hispanic Maya Culture, Objects, and Place at Chau Hiix, Belize. PhD dissertation, Department of Anthropology, Indiana University, Bloomington.
- Wrobel, Gabriel D. G.
2003 Metric and Nonmetric Dental Variation Among the Ancient Maya of Northern Belize. PhD dissertation, Department of Anthropology, Indiana University, Bloomington.

AUTHOR INFORMATION

Caroline Beebe ■ Affiliate, Department of Anthropology, Indiana University, Bloomington, Indiana 47405, USA (carolinebeebe@hotmail.com)