Digitally-Mediated Practices of Geospatial Archaeological Data: Transformation, Integration, & Interpretation

Heather Richards-Rissetto
University of Nebraska-Lincoln, richards-rissetto@unl.edu

Kristin Landau
Alma College, landaukv@alma.edu

Follow this and additional works at: https://digitalcommons.unl.edu/anthropologyfacpub

Part of the Archaeological Anthropology Commons, Digital Humanities Commons, Geographic Information Sciences Commons, Landscape Architecture Commons, Other History Commons, Other History of Art, Architecture, and Archaeology Commons, Social and Cultural Anthropology Commons, Spatial Science Commons, and the Urban, Community and Regional Planning Commons

https://digitalcommons.unl.edu/anthropologyfacpub/163

This Article is brought to you for free and open access by the Anthropology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Anthropology Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Digitally-mediated practices of archaeological data require reflexive thinking about where archaeology stands as a discipline in regard to the ‘digital,’ and where we want to go. To move toward this goal, we advocate a historical approach that emphasizes contextual source-side criticism and data intimacy—scrutinizing maps and 3D data as we do artifacts by analyzing position, form, material and context of analog and digital sources. Applying this approach, we reflect on what we have learned from processes of digitally-mediated data. We ask: What can we learn as we convert analog data to digital data? And, how does digital data transformation impact the chain of archaeological practice? Primary, or raw data, are produced using various technologies ranging from Global Navigation Satellite System (GNSS)/Global Positioning System (GPS), LiDAR, digital photography, and ground penetrating radar, to digitization, typically using a flat-bed scanner to transform analog data such as old field notes, photographs, or drawings into digital data. However, archaeologists not only collect primary data, we also make substantial time investments to create derived data such as maps, 3D models, or statistics via post-processing and analysis. While analog data is typically static, digital data is more dynamic, creating fundamental differences in digitally-mediated archaeological practice. To address some issues embedded in this process, we describe the lessons we have learned from translating analog to digital geospatial data—discussing what is lost and what is gained in translation, and then applying what we have learned to provide concrete insights to archaeological practice.

Keywords: Digitally-mediated archaeology; geospatial; archaeological practice; historical approach; Mesoamerica; data intimacy; paradata
(Allison 2008; Gaffney, Stančič & Watson 1995; Smith 1995; Wylie 2017). As archaeologists, we collect primary, or ‘raw’ data of extant archaeological features and artifacts that we often use to create ‘derived’ data such as maps, 3D models, and statistics based on post-processing, analysis, and interpretation (Beale and Reilly 2017; Costa et al. 2013; Faniel et al. 2013; Huggett 2015; Kansas and Kansas 2018; Kintigh et al. 2017). These raw data often need to be digitized—changed from analog to digital—to be useful for digital technologies and methods.

Archaeologists employ numerous software including databases, Geographic Information Systems (GIS), photogrammetric tools, and 3D environments to process, integrate, and analyze archaeological data; in other words, we transform analog and digital ‘items’ to create new (derived) data for archaeological research. Additionally, we create ‘natively digital,’ ‘digital-first,’ ‘digital-exclusive,’ or ‘intrinsic born’ digital data (Austin 2014). In contrast to digitization of analog data, such natively digital data come from post-processing primary data or generating data that do not or did not have a physical counterpart. This process creates new challenges and opportunities in archaeological scholarship (Digital Preservation Coalition 2015; Forte 2014). For example, we use flatbed scanners to digitize a site map recorded in a field notebook to convert the analog page into a digital format such as a Tagged Image File Format (TIFF). While a TIFF is machine-readable, the data require post-processing to be useful for geospatial analysis. For example, the scanned map must be georeferenced to provide real-world coordinates and scale; it must also be vectorized to provide data for analysis in a GIS or other platform. In other words, post-processing necessitates multiple steps of human decision-making, producing numerous file types, and results in new data. The creation of new data through digitization, most often but not always through post-processing, is called datafication.

Some scholars define datafication as “transforming objects, processes, etc. in a quantified format so they can be tabulated and analysed” (Gattiglia 2015: 115, emphasis ours; Mayer-Schönberger and Cukier 2013). In contrast, we contend that datafication outputs are not limited to quantifiable data, and we recommend shifting the definition of datafication to emphasize process i.e. the transformation and translation of objects and processes rather than outputs (Richards-Rissetto 2017b). Basically, in contrast to digitization which ‘replicates’ original data, datafication creates derived, or new data, which requires human translation (interpretation) and encourages unique considerations for archaeological scholarship. Datafication of both born-digital and analog formats offer archaeology more than either can do alone.

Datafication involves what digital scholars call metadata and paradata. While the term metadata describes information about the data themselves (Clarke 2015; Esteva et al. 2010; Hodder 1997; Roosevelt et al. 2017; Ullah 2015; Witcher 2008), paradata more specifically concern interpretive decisions. Recording paradata, the “information choices or the process of interpretation so that the aims, contexts and reliability” of methods can be evaluated (Bentkowska-Kafel and Denard 2012:1) is a major challenge. With digital data, in particular geospatial and 3D modeling and visualization, archaeologists can easily modify raw and derived data to generate new derived data. However, retracing our steps is not straightforward, though transparency is necessary for others to assess data quality as well as analytical results (Huggett 2014; Kansas et al. 2010). Datafication mandates not only metadata but also paradata, thus requiring unique practices for digital scholarship in archaeology (see below). However, datafication also brings a great opportunity for data intimacy: a deep familiarity with the data that affects perception and affords new insights (Cavillo and Garnett 2019; Fahmie and Hanley 2008; Hong 2016). Intimacy is increasingly essential for a digitally-mediated archaeology in which data transformation, integration, and creation is anything but straightforward.

Archaeological data is heterogeneous, making not only the data messy, but perhaps more importantly making the scientific process itself messy; research does not proceed in an orderly series of steps (Boyer 1990). Yet this messiness affords new opportunities for data integration that require deep interdisciplinary thinking and often lead to innovative methods and analyses (Demjän and Dreslerova 2016; Harrison 2018; Kansas 2010; Kintigh 2006; von Schwerin, Lyons et al. 2016). Even in cases where digitization/datafication standards or best practices exist (e.g. Open Geospatial Consortium), researchers still must make numerous decisions as we generate digital data. This decision-making process is not a new aspect of digital archaeological practice (Hodder 1997; Hodder 2003). For example, in hand-drawing profiles we decide on important points (x, y, and z) to map based on previous knowledge, experience, objectives, etc. However, in generating born-digital and derived digital data we often make black box decisions (Caraher 2016) based on convention or ‘mysterious’ software algorithms. Given the emerging nature of digital technologies and tools, we often make decisions based on trial and error, searching the internet for solutions, or contacting colleagues. Also, because digital data are dynamic (Alberts, Went & Jansma 2017), our initial choices more easily and readily change, leading to new challenges and advantages in digitally-mediated archaeology.

Because of the dynamic nature of digital data and technologies, we contend that digitally-mediated data transformation, integration, and interpretation require reflexive, iterative thinking—we must be more aware of our decision-making processes (Engel and Grossner 2014; Esteva et al. 2010; Hodder 1997; Hodder 2000; Hodder 2003; Lukas, Engel & Mazzucato 2018; Roosevelt et al. 2015; Tringham 2010). Why and how do we make specific choices? And how can we document our choices, i.e. record metadata and paradata, to allow for digitally-mediated scholarship to become better integrated and accepted into archaeological practice? These questions are part of larger challenges and opportunities of digital scholarship in archaeology.
In the first part of this paper we introduce a historical approach for digitally-mediated archaeology. Derived from interdisciplinary collaboration between historians and historical archaeologists, the approach encourages increased reflexivity and critical analysis of data sources. Just as archaeologists study the position, form, material, and context of an artifact, historians consider the same in scrutinizing documentary resources. Here, we explore how to apply source-side criticism of (sometimes actually historical) analog and digital data.

In the second part of this paper, we discuss several examples of translating analog data to geospatial digital data including: (1) converting maps originally generated with alidade and plane table to Geographic Information Systems (GIS) data, (2) converting hand-written field notes into GIS data, (3) integrating multi-source data (i.e. vectorized maps, GNSS, total station, and airborne LiDAR), (4) processing data to generate georeferenced 3D models, and (5) analyzing digital data in different software for scholarly research and interpretation. We summarize the lessons we have learned from our experiences in transforming analog data to geospatial digital data, discussing what is lost and what is gained in translation, and then applying what we have learned to provide concrete insights to archaeological practice. We contend that as we transform, integrate, and analyze these data, we are not simply digitizing data but rather we are performing datification. In other words, we are acquiring new knowledge about data collection, documentation, processing, and interpretation, which can lead to new archaeological questions and methodologies and enhance the nature of archaeological scholarship

Archaeological scholarship, whether digital or not, stems from specific research goals. We ask questions that guide our research design from data collection to analysis to dissemination. To situate our discussion, we use a case study from the ancient Maya site of Copán, Honduras, that has specific research goals and questions related to landscape archaeology using a combination of analog and digital data.

2. A Historical Approach for Digitally-mediated Archaeology

In anthropology, the so-called postmodern turn of the 1980s encouraged greater awareness of power differentials between observed and observer. The concept of culture itself was scrutinized as a reification and tool for “othering” (Abu-Lughod 1991; Clifford and Marcus 1986). Anthropologists began to question themselves, the ethnographies they produced, and the epistemological basis for the scientific hypothetico-deductive-nomological approach. Customs, traditions, and ways of being could only be understood within their appropriate contexts. Given that anthropologists can never actually get inside an informants’ head, postmodernists argued that their books are merely one-sided accounts or fictions, and so they should be treated just as any other literary text (e.g., Salzman 2002). In this vein, post-processual archaeologists attempted to “read the past” or “read material culture” as a way to construct meaning (Hodder 1984; Tilley 1990; Tilley 1993).

Because historical archaeology involves both artifacts and texts—material objects as well as writing—postmodern critiques had much to offer. Until then, while written histories provided a more privileged position than archaeological data, they were not subjected to the same kinds of rigorous contextual analyses as artifactual studies (Lightfoot 1995; Morrison and Lycett 1997; Stahl 1993). Historical archaeologists began treating texts as artifacts by more reflectively considering their contexts, how they obtained them (source-side criticism) and how they applied them (subject-side criticism). Of particular importance was source-side criticism, and archaeologists followed the lead of historians in more carefully assessing the authenticity and validity of documentary accounts. W. Raymond Wood (1990) argued that archaeological records, photographs, maps, and the landscape itself be considered ‘documents,’ and thus open to the same kind of source-side criticism as historical texts. He summarizes the historical method in four steps: (1) formulating the problem or research question for which documents are needed, (2) determining which documents are authentic (‘external criticism’), (3) determining which details within a document are credible (‘internal criticism’), and (4) organizing all reliable information into a narrative to resolve the research problem.

Wood’s (1990) first step of formulating a research question is encapsulated by archaeology’s turn to ‘problem-based research’ during 1960s processualism, and later on in GIS. For example, Lock and Stančič (1995: xiv) stressed that it is not the specific mathematical procedures themselves that will be the future of innovative research with GIS but rather “the underlying archaeological approaches and questions determining their use.” Thus, as in dirt archaeology, a preconfigured research question is also a necessary starting point for digital archaeology. Wood’s (1990) second step of external criticism involves assessing a document itself, while his third step, internal criticism, addresses the document’s specific contents and meaning. To perform external criticism, one must focus on the author and date and obtain the original version rather than a copy. Next, the researcher must separate the content of the document into eyewitness accounts at the moment versus descriptions written by another individual or later. Most important in evaluating credibility is temporal proximity to the event; next is consideration of potential distortion due to the intended purpose and audience of the document; last involves addressing the competency and expertise of the writer and whether there is independent corroboration (Wood 1990, 89). Also important for our purposes is Wood’s admonition to
We examine five types of data transformation that have historical method to geospatial data in archaeology. We apply the above insights on analog/static versus digital/dynamic data and source-side criticism to text, maps, and 3D models.

3. Geospatial Data: What’s the Big Deal?
Archaeology is all about location. Provenience is essential across scales. The more we know about location, the greater potential for more informed and granular interpretations. Archaeologists began to employ GIS originally for data management and not long after for spatial analysis (Connolly and Lake 2006; Wheatley and Gillings 2002). GIS revolutionized the way archaeologists deal with spatial data, and the question as to the magnitude of its impact on archaeological theory is still debated (Howey et al. 2013; Gunnarsson 2018; Tringham 2010; Wells et al. 2001, Llobera 2007a, Llobera 2007b; Richards-Rissetto 2017b). Nevertheless, GIS brought greater awareness to the potential of geospatial data for archaeological studies. No longer would our maps be aligned to site-scale coordinate systems based solely on an arbitrary (0, 0) origin. Rather, our site data could be tied to real-world coordinates allowing us to overlay multiple layers of data such as geology, geomorphology, and land cover, and importantly for landscape archaeology, tied to a much larger area with greater precision allowing new types of analyses.

Today, many archaeologists have GNSS to acquire data points for not only site location but millimeter-level geospatial data of intra-site features through integrating a variety of digital tools (e.g. total station, laser scanning, and photogrammetry). Others have legacy data from earlier surveys, excavations, and analysis (Allison 2008; Clarke 2015; Faniel et al. 2013; Kansa and Kansa 2018; Ullah 2015; Witcher 2008), which provide data that are ‘lost’ due to the destructive nature of excavation, urbanization, agriculture, looting, taphonomic processes, natural disasters, and more (e.g. Gruen, Remondino & Zhang 2004). These analog data provide a rich source of information that can be converted to and subsequently integrated with digital data to generate not only new data, but to lead to new forms of archaeological practice and scholarship (Faniel et al. 2013; Gunnarsson 2018; Tringham 2010; Wells et al. 2014; Wylie 2017). The use of digital geospatial data has revolutionized the practice of archaeology, but archaeologists must still be vigilant of its origins and context (Ullah 2015).

4. Methods, Lessons, & Reflections: Translating Analog Data to Geospatial Digital Data
We apply the above insights on analog/static versus digital/dynamic data and source-side criticism of the historical method to geospatial data in archaeology. We examine five types of data transformation that have been particularly relevant to our own research, and provide a few words on the experience as well as lessons for future practitioners. We illustrate the five transformation types with different categories of data from the ancient city of Copán. Before outlining the five types of data transformation, we provide a brief history on the kinds of analog and digital data that currently exist for Copán.

The ancient Maya site of Copán has a long occupation dating back to at least 1800 BCE. Today, it is a UNESCO World Heritage Site in Honduras, but from the fifth to ninth centuries it was the seat of a dynastic kingdom that at its peak governed over 250 square kilometers (Bell et al. 2004; Fash 2001). Excavation dates back to 1834 when Guatemala’s governor, Juan Galindo, mapped part of the site’s core and excavated a tomb in the main civic-ceremonial complex (Fash and Agurcia Fasquelle 1996). Unfortunately these primary data are lost, but in 1869 Stephens and Catherwood—two early explorers of Central America—described, mapped, and created drawings (using a camara lucida) of Copán’s jungle-covered main civic-ceremonial core (Stephens and Catherwood 1841). In the early to mid-nineteenth century, archaeologists began scientific studies of the site that included excavation, architectural drawings, and maps (Maudslay 1889–1902; Morley 1920). Later, in the late 1970s and early 1980s archaeologists carried out a 100% pedestrian and mapping survey of 24 square kilometers surrounding Copán’s main civic-ceremonial complex (Fash and Long 1983). In the early 1980s two Austrian architects used photogrammetric methods to generate large-scale (1:200) maps of the main civic-ceremonial complex (Hohmann and Vogrin 1982). Additionally, maps from individual excavations are available via unpublished field notes, type-written summaries of field notes with penciled-in additions, type-written finalized reports, dissertations, monographs, and other publications available online and in Copán’s onsite archives. These maps along with archival and published data provide a wealth of analog resources to investigate ancient Copán.

In following the first step of Wood’s (1990) historical method, we want to be explicit in defining the nature of the problem for which we seek documentary sources. Generally speaking, our case study has two broad research questions: (1) What is the nature of social interaction at Copán in the late eighth to early ninth centuries, just prior to the city’s decline? and (2) How did daily life within Copán’s urban neighborhoods change over time in relation to major political and/or economic events? To examine these questions, we focus on accessibility and visibility within the city of Copán. We ask: who lived in view of royal architecture? Who was visually isolated? Were certain social groups channeled toward specific locations? If so, for what purposes? Additionally, can measures of accessibility and visibility provide data useful for identifying neighborhood or other boundaries (Landau 2015; Llobera, Fàbrega-Álvarez & Parcero-Oubiña 2011; Llobera 2001, Llobera 2007a, Llobera 2007b; Richards-Rissetto 2010)?

In order to address these questions, we need not only geospatial data, but multiple scales of data from
excavation units to regional surveys, and many of these data are only available in analog formats. Thus, we developed the above mentioned five-step process that includes: (1) converting maps—originally generated with alidade and plane table—to GIS data, (2) translating hand-written field notes into GIS data, (3) integrating multi-source geospatial data (e.g. digitized analog data with GNSS data, total station, and airborne LiDAR data), (4) processing GIS and other data to generate georeferenced 3D models, and (5) analyzing geospatial digital data in different software for scholarly research and interpretation.

4.1. Step 1: Digitizing, georeferencing, & attributing paper maps

Lessons: While labor-intensive and time-consuming, the process of digitizing, georeferencing, and attributing maps created with differing methods, at multiple scales, and in different languages (English, Spanish, German, and French), and then painstakingly vectorizing them provided new insights and sparked new archaeological questions about the Mahler, or prismatic, method of mapping Maya sites (Hutson 2012) and Copán’s site typology (Richards-Rissetto 2010, 2012; Richards-Rissetto and Landau 2014; Willey and Leventhal 1979). Our experience was similar to that of Ullah’s (2015) exploration of the minute details and large errors of legacy survey data in Jordan. Here Wood’s (1990) second and third steps apply: the various paper maps must be subjected to external criticism (are they authentic?) as well as internal criticism (are the details within accurate?). Making such judgments inherently involves becoming well acquainted with the context of creation for each map: what we term data intimacy. What were standard cartographic practices in the US, Honduras, Germany, and France in the 1970s? What were the defined problems for which these maps were produced? Which details did the mapmakers include and exclude, and why? Do field notes admit to mistakes, illnesses, land-access issues, etc. that were or were not published in the final map? Did individual field workers have years of experience, or did they learn on the job? Such contextual questions must be considered in the digitization process. Reflexively as well, the digitizer should explicitly record which maps (external criticism) and which details (internal criticism) were actually digitized or left out, and why (Clarke 2015; Esteva et al. 2010; Hodder 1997; Roosevelt et al. 2017; Ullah 2015; Witcher 2008). While digital implies speed—archaeologists quickly acquire millions of 3D points using a laser scanner—we learned that the best practice is slow practice (Caraher 2016): to take a step back and critically consider the longer term implications of digitization before jumping in. It is critical to examine all mapped analog (and digital) data before georeferencing and vectorizing. Careful examination may help to identify an appropriate grid system and lay out a methodology suited to heterogeneous data (Demján and Dreslerova 2016).

In the case study, maps ranged from twenty-four 1 km square plane table and alidade maps at a scale of 1:2000 (Figure 1) (Fash and Long 1983) to 1:200 scale photogrammetric maps of Copán’s civic-ceremonial core (Hohmann and Hohmann-Vogrin 1982), to excavation maps of individual sites (Maca 2002; Webster 1989). After researching how each of the existing maps were created, we decided to georeference them to the Copán Archaeological Project (PAC 1) site grid (Fash and Long 1983) for several reasons. First, it offered the best tie points for Copán’s heterogeneous mapped data. It also provided a way to double-check and link attribution because structure and group names are based on grid quadrants with additional data (e.g. site type, number of plazas) available in a separate volume (Fash and Long 1983). Third, the Copán site archives contains a massive collection of original field notes, type-written versions of the field notes, original hand-drawn maps, reports to funding agencies, and final publication drafts. Such sources were helpful in determining how much to rely on particular internal details. For example, when an archaeologist’s field notes indicated that an area was heavily forested, we noted that archaeological structures and contour lines may be less accurate here than in other areas.

Figure 1: Example of 1:2000 scale plane table and alidade map, from Fash and Long (1983) (left) and photogrammetric map at scale 1:200, from Hohmann and Vogrin (1982) (right).
Due to fine lines, no color differentiation, and multiple data layers in a single map (e.g. hydrology, structures, contours, modern roads, text), we manually digitized the maps to create georeferenced vector data (i.e. shapefiles) to ensure accurate data capture (Richards-Rissetto 2010). Three data layers were vectorized—contour lines, archaeological structures, and hydrology—and attributed with Group name, Structure name, Site Type, and Elevation using data from maps, architectural drawings, and text. Circling back to Wood (1990), it is essential in digital archaeology to capture not only metadata, but paradata (Bentkowska-Kafel, Denard & Baker 2012; Denard 2012); that is, recording the data sources, methods, etc. that inform the choices we make as we digitize, and importantly providing information on any modified data, for example, filling in missing gaps on a map using excavation data or architectural drawings. Capturing metadata and paradata is essential in digital archaeology to allow other researchers to reproduce not simply an end-product, but to actually retrace our processes to verify as well as build on such scholarship. In the end, such practice will also facilitate data preservation and access and help formulate best practices and standards because it allows data to be readily re-used (Richards-Rissetto and von Schwerin 2017).

Another advantage of manual digitization is data intimacy. On-screen tracing of archaeological features by hand simulates traditional hand-drawn mapping practices. Such a process provides familiarity with second-hand data that is lost with automatic vectorization. For example, in the 1970s, cartographers created a five-level typology for classifying aboveground architectural remains (Fash and Long 1983; Willey and Leventhal 1979) that archaeologists adopted to represent socioeconomic status. Through manually digitizing and attributing over 3000 structures, we came to question the validity of correlating the typology to social status (Richards-Rissetto 2010). Our suspicions were later supported because there was no spatially statistically significant difference in accessibility between some elite (type 3) and non-elite (type 2) residential groups (Richards-Rissetto 2012; Richards-Rissetto and Landau 2014). Therefore, the slow and tedious practice of on-screen tracing led to the development of a research question about accessibility between people of different socioeconomic status. Results from this study led to a correction in the Copán site typology, changing our understanding of the nature of status differences and inequality at the ancient city. Applying Wood’s historical method encouraged new research questions that ultimately helped us better answer major anthropological questions.

Basic lesson: Although today’s digital archaeology allows rapid and efficient digitization and datafication, we should step back and slow down. Our experiences have shown that developing data intimacy—though sometimes hours of painstaking manual digitization—affords greater exploration and reflection on the data. Gaining introspective clarity during the process of digitization and datafication may lead to significant new research findings, previously unconsidered.

4.2. Step 2: Translating archival documents into spatial data & informing the geospatial process

Lessons: Archival field reports and hand-written notes are an often untapped resource; however, such data are inconsistent—some investigators write more than others and notes are missing, often unstandardized (between individuals, between projects, and across time), and provenience data are hit or miss. Moreover, documents are composed in multiple languages, and at times the writing is illegible (e.g. Clarke 2015, Ullah 2015, Witcher 2008). Nonetheless, after applying Wood’s (1990) criteria to determine credibility, these archival data are worth the effort—they fill in missing pieces and enrich research. For example, they provide attributes for mapped features, rationale for terminology and methods, and ‘lost’ provenience.

In the case study, we scanned archival data from the library at the Center for Regional Archaeological Investigations (CRIA) at Copán Ruinas in Honduras. These data include hand-drawn maps and profiles, artifact counts, provenience information, catalog numbers, etc. We scanned the originals as PDF files (for documents) and TIFF files (for images and maps) to address three interrelated goals: for long-term archival purposes, to assist the CRIA in digitization efforts, and to gather more information on precisely how archaeological structures were interpreted and mapped. While we were reasonably sure that all field notes and reports were authentic due to their curation at the site archives, we combed through these documents for spatial information that we could transform into usable geospatial data. We read each source completely to gain a sense of internal validity—does the author contradict themselves? Are peculiar margin comments corroborated by other authors within the archives? Once we established validity and accuracy, we georeferenced and vectorized maps into shapefiles, and populated spreadsheets with attributes linked to the shapefiles.

A key challenge was to assign height to the archaeological structures. In the Maya region, surveyors record the length, width, and height of architectural mounds (i.e. collapsed structures), but do not estimate original structure heights. To estimate structure heights we began by gathering spatial and other relevant data from archived excavation notes, published monographs, and ethnographic data. In particular, annotations and their placement within archival documents provided insights (often lost in the typewritten field reports) via rough sketches and from architectural materials and construction techniques (Figure 2) (Tringham 2010). Beyond providing X, Y, and Z spatial data, these data were integral to developing a GIS method to estimate height based on site type, construction materials, and excavation data (Richards-Rissetto 2010, 2013). Ultimately we estimated height using a trigonometric function, but developing mathematical formulas and an appropriate methodology required a close reading of various analog sources.

Basic lesson: Texts should be treated as artifacts themselves (Lightfoot 1995; Morrison and Lycett 1997; Stahl 1993); we cannot simply take them as fact and incorporate
them, but rather we need to try to understand the context and purpose of their writing in relation to the writer and historical circumstances. Each document has its own historical trajectory and materiality.

4.3. Step 3: Integrating multi-source geospatial data (e.g. Shapefiles, GPS, GNSS, Total Station, LiDAR)

Lessons: Combining different geospatial datasets typically fills gaps in archaeological maps, giving a more complete picture despite differences in original acquisition or granularity. However, sometimes different datasets overlap. How do we decide which dataset is best, how to combine datasets, or how to give more weight to the ‘better’ dataset? The second and third step of Wood’s (1990) historical method (external and internal criticism) again become important in integrating various datasets. First, how was each dataset initially produced? For which research questions were the data commissioned to answer? Second, which aspects of each map were more ‘accurate’ in instances of overlap? We conclude that deciding which representation is more ‘correct’ or ‘accurate’ should be an iterative process and ideally best accomplished while in the field, where ground-checking is possible. No one type of data (capture) is necessarily ‘better’ than others, but rather each data type comprises parts—some parts are more useful or accurate than others.

In our case study, total station-based mapping in San Lucas revealed what appeared to be a ‘new’ archaeological group—unmapped in previously published reports. We also ‘lost’ a group that had been previously mapped, which we could not relocate on the ground (similar to Ullah’s [2015] experience). Consulting LiDAR data showed that the originally mapped group had been erroneously placed. While the internal architecture was mapped correctly, the group was placed about 200 m away from its actual location. Therefore, these two groups were one in the same. In another example, while total station data captured low mounds (Landau, Richards-Rissetto & Wolf 2014), it was difficult to differentiate low archaeological mounds (<25cm) from natural topography using airborne LiDAR (von Schwerin et al. 2016). In the process of integrating multi-source datasets, we learned that datasets can ‘self-correct,’ but only if we iterate back and forth between them to reveal which bits are more or less accurate. In the end, we create a critical combination of all maps—by applying Wood’s method—that results in improved accuracy and precision all around.

Figure 3 is an example from the neighborhood of San Lucas at Copán (Landau 2016). It illustrates overlaid data gathered from three different sources—pink (Fash and Long 1983), yellow with black lines (Landau 2016), and a LiDAR-derived landscape (von Schwerin et al. 2016). The Fash and Long (1983) data were collected using alidade and plane table at a time when the Copán Valley was much more sparsely occupied, and this area was likely a cow pasture with low to medium overgrowth. Wolf and Landau re-mapped this architectural group in 2012–14 with several different GNSS units and a total station with prism, and in 2013, the MayaArch3D Project commissioned LiDAR data (von Schwerin, Richards-Rissetto et al. 2016). In general, Wolf and Landau consulted the Fash and Long (1983) maps while in the field using GNSS...
receivers and total station with a prism. First we searched for the structures as indicated on the 1983 maps. Keeping these structures in mind with the contemporary landscape topography, Wolf drew the architectural group as he interpreted it by hand in a notebook; afterward we took a series of three to six points for each structure. Wolf later reconciled these points with his hand-drawn maps (see Figure 2). Afterward, when plotting the 1983 maps together with Wolf’s maps on top of the LiDAR hillshade surface, Landau made further corrections to the Wolf drawing. For example, she modified the edge of the flattened area in the northwest corner of Figure 3, to give a more accurate sense of its extent.

Another lesson involves careful, critical use of automatic digitization tools. While the vector to raster tool in GIS is push-button (not quite black box, but easily non-critically applied), dealing with architecture rather than topography requires different decisions, methods, and tools. For example, what spatial resolution is sufficient? To capture details such as platforms and stairs require high-resolution data; however, generating a 10 cm raster surface for 24 square kilometers or more requires high levels of processing power—often not available to individual researchers or archaeologists in developing countries. Additionally, in cases of landscape analysis, these rasterized architectural data also need to be integrated with the terrain (topographic surface). While LiDAR data are available in some areas, typically they are still unavailable to archaeologists due to high costs and lack of flights, particularly in remote regions. Thus, our options for free or low-cost raster terrain data are limited to lower-resolution datasets such as Shuttle Radar Topography Mission (SRTM) or Advanced Spaceborne Thermal Emission and Refraction Radiometer (ASTER), which unfortunately are not sufficient for visibility analyses within urban landscapes such as ancient Maya cities where topography is integral to site layout (Aveni and Hartung 1986; Gagnon et al. 2011; Inomata 2008; Juarez, Salgado-Flores & Hernández 2019; Landau 2015; Richards-Rissetto and Landau 2014). Another option is analog data acquired via instrument mapping and published as paper maps with contour lines. These paper maps typically provide a larger-scale (i.e. higher resolution) terrain than free DEM data (particularly outside of the U.S. and Europe), but following a historical approach, we must step back to critically evaluate the quality of source data.
Basic Lesson: All sources provide complementary and unique information that together create a more holistic and empirical picture. Digitally-mediated practice necessitates and allows us to more deeply interrogate data accuracy and interpretation, particularly through an iterative process. Importantly for archaeological practices, geospatial data integration raises questions such as: Can we identify spatial patterns by examining similarities and differences among analog maps, LiDAR, and excavation data? Can such comparisons help us interpolate older analog maps? How accurate are LiDAR data in particular cultural and environmental contexts? Can we devise algorithms to more accurately detect low mounds by ground-checking a stratified sample and comparing topography and vegetation to algorithm-detection accuracy? These questions impact archaeological practice and digital scholarship.

4.4. Step 4: Processing data to generate georeferenced 3D models

Lessons: In the past decade, particularly since the advent of out-of-the-box photogrammetry (i.e. Structure from Motion), 3D data have become commonplace in archaeology. However, most 3D data are not born-digital, but rather they are acquired in the field, lab, or museum capturing physical objects and landscapes. These primary data can be instantly georeferenced, or not, depending on available technology and the location of data capture. However, converted analog data such as structure maps introduce new challenges as we move from 2D (vector) to 2.5D (raster) to 3D models (mesh/faces). While we can transform analog maps to GIS vector data and subsequent raster data, our 3D results are extruded schematic models lacking (slanted) roofs, architectural sculpture, and often platforms and stairs depending on the original map. Transforming 2.5D data into true 3D models usually necessitates manual modeling, though procedural modeling is now offering innovative opportunities (Saldana 2015).

While directly generating 3D architectural models from GIS (2.5D) data is not ideal, it offers the benefits of conveying uncertainty and offering a close-reading of data. 3D models, particularly those that are photo-realistic, can lead viewers to false certainty about reconstructions (Kantner 2000). However, abstract models (perhaps augmented by transparency or color-coding) portray important ambiguities (Brunke 2018; Kensek, Dodd & Cipolla 2004; Lengyel and Toulouse 2015). Considering Wood’s (1990) process in reverse, how can we use 3D modeling to indicate instances of uncertainty regarding source authenticity and accuracy? Creating data that includes measures of uncertainty would allow future researchers to more easily apply source-side criticism and, ultimately, correction. Moreover, manual 3D modeling leads to data intimacy providing new insights. For example, ambiguities in mapping, typically not identified in procedural modeling, can be identified and then employed to write scripts to generate empirically-informed procedural models. Yet, we still end up with static, fixed models that represent a single interpretation (i.e. reconstruction). In this scenario, we fail to take advantage of certain digital affordances; that is, we do not take advantage of digital technologies to generate multiple hypothetical 3D models (or simulations) for structures or landscapes.

Thus, in the case study we turned to procedural modeling, i.e. ruled-based rapid generation of buildings from GIS data (Richards-Rissetto and Plessing 2015), to generate multiple simulations. We generated 3D models from a spatial database with metadata and the decisions we made (i.e. paradata) stored both as a text document and schematic hierarchy—offering innovative possibilities for digital data storage, accessibility, and reuse (Bentkowska-Kafel, Denard & Baker 2012; Denard 2012; Esteva et al. 2010; Faniel et al. 2013; Lukas, Engel & Mazzucato 2018; Richards-Rissetto and von Schwerin 2017). Additionally, these procedural models provide information to digitally define basic elements and components of ancient Maya architecture, which scholars have sought to define for over one-hundred years (Andrews 1997; Kubler 1905). Using architectural definitions (Loten and Pendergast 1984), we created rules for elements and components that allow for dynamic modeling rather than static modeling of architecture—this digitally-mediated process facilitates hypothesis generation with empirical underpinnings that are documented in procedural modeling scripts (Richards-Rissetto and Plessing 2015).

Basic Lesson: In part because of the time input for manual modeling, singular 3D architectural models can mislead viewers to false impressions of the past. Procedural modeling of geospatial data into 3D introduces new possibilities because we can create multiple simulations based on different data sources. Given that each model displays a different set of conclusions based on the data—and, importantly, includes the source data on which that particular conclusion was based—procedural modeling provides more dynamism to archaeological data. This allows researchers to evaluate multiple different scenarios, and could potentially reveal to the public the complexities of digital 3D archaeological reconstruction.

4.5. Step 5: Analyzing digital data in different software for scholarly research & interpretation

Lessons: While ‘analysis’ occurs in data translation, GIS, 3D modeling software, and VR afford opportunities for knowledge generation via integration, computation, and visualization (Forté and Pescarin 2012; Jones and Levy 2014). Each software offers unique tools and methods that facilitate, enhance, and ultimately change archaeological practice and scholarship. Yet through reflectively iterating between these software, we afford additional new possibilities. While GIS provides tools to convert analog data to geospatial digital formats, its power for scholarship resides in its analytical capabilities. Using GIS we can identify spatio-temporal patterns and trends of big and complex data to investigate old questions and hypotheses in alternative ways and propose new lines of inquiry. In the case study, using GIS we developed computational visibility and accessibility approaches across multiple scales to investigate social connectivity among Copán’s different socio-economic groups (Landau 2015; Richards-Rissetto...
The process of creating 3D models was not linear but rather we iteratively worked back and forth among GIS, LiDAR, and excavation data necessitating a deep exploration of the data as parts but also as a whole. This data intimacy led to new questions about the Mahler method of mapping ancient Maya sites, which records mound heights and not actual structure heights, and thus proves problematic for direct GIS to 3D model conversion. Additionally, Copán’s Site Typology attributes sites from Types 1-5; however, site types refer to the ‘highest’ socioeconomic status of the entire group and do not provide information on lower-status occupants or on structure functionality—both of which affect 3D modeling and subsequent archaeological interpretations.

GIS and 3D models (reality-based and reconstructions) provide source data to create 3D virtual environments of ancient Copán using, for example, VR and procedural modeling. However, other, originally analog, data also provide essential information to create 3D simulations of past landscapes that serve as more than pretty illustrations. They enable us to create multiple simulations to interchange data, investigate old hypotheses, and create new interpretations. In these simulations, analog data are just as essential as digital data because they provide information on features that are now lost to degradation, urbanization, excavation, or other processes. Architectural hand-drawings, archival photos, and field notes fill in data gaps. As we go from analog to digital and subsequently integrate datasets, we do not simply convert data, but we translate it—we see anomalies, errors in data, find ‘missing data,’ and think about typologies or classification schemes. In other words, we acquire data intimacy. With these 3D simulations, we have the ability to convey data ambiguity (Brunke 2018; Kantner 2000; Kensek, Dodd & Cipolla 2004), explore our data in unique, dynamic, and experiential or embodied ways (Forte and Pescarin 2012; Forte and Pietroni 2009; Richards-Rissetto et al. 2012; Richards-Rissetto et al. 2013), and perform landscape-scale analyses that are impossible without going digital.

**Basic Lesson:** In the process of translating analog data to digital form, various technologies including GIS, 3D modeling, and VR offer new pathways for data integration, computation, and visualization. Although GIS provides a suite of analytical tools, its 2.5D format prevents crucial architectural and landscape features from playing their part in visibility studies, for example. Therefore 3D modeling and VR take over where GIS leaves off: the procedural modeling process is predicated on a back-and-forth agreement and decision-making among all datasets, static analog and dynamic digital. The product is more than just a pretty picture because it can show multiple possibilities, re-open preliminary conclusions, and close lasting questions.

**5. Discussion—Lessons Learned in Transforming Analog to Geospatial Digital Data**

Our particular experience working with geospatial data at an archaeological site with over 100 years of excavation history necessitated translating various analog data into digital data. Because we are dealing with raw and derived geospatial data, the steps of the digitization and datafication process are complex; therefore, we aimed to provide some perspective to help guide others in an area for which standards and best practices are emerging. The historical approach we advocate (following Wood 1990) provides a methodology for assessing the origins and accuracy of static analog and born-digital data. In converting different  

Figure 4: GIS Map of Group 12M-1 from San Lucas neighborhood, Copán (left); 3D SketchUp reconstruction of Group 12M-1 (right).
digitally-mediated practices of geospatial archaeological data. We advocate a historical approach to digitally-mediated data translation, as well as the iterative process between all data toward a more holistic, more accurate representation.

In the creation of 3D virtual environments from 2 or 2.5D data, using procedural modeling and VR allows archaeologists to interrogate and integrate various datasets. Through the process of transforming disparate datasets such as architectural drawings, excavation notes, and archival photos into useful digital data that forms part of the 3D simulations, we develop data intimacy—identifying key pieces of information that would be lost in automatic methods that simply convert data rather than translate data as required by a close reading. Therefore, as we translate analog to digital data, we develop a deeper understanding and appreciation for how the data that are digitized came to be. The digital affords archaeologists greater intimacy with both legacy datasets (analog and digital) as well as derived data and the intermediate files created through datafication. Several of our examples above demonstrate the intellectual advantages of data intimacy and slow science (sensu Caraher 2016).

We learn about the archaeology we are studying through the act of ‘translating’ these data. Importantly, we invite similar reflection of already digital data because often we have already lost some of the history of these data, especially before metadata or paradata were emphasized for inclusion. Digital data can give a false sense of accuracy because they are often clean and ready-to-use; likewise, while digital data allow landscape-scale analyses that are impossible with analog formats, they impart a distance, disembodied, and masculine god’s eye view. In both cases, the downside is that we often forget the palimpsest from which the data originally derived, as well as the time, material conditions, labor, and small decisions that went into collecting them. In a sense, we experience another black box stemming not only from unknown or poorly understood algorithms, but also from a process that generates intermediate data sets, i.e. datafication, and datadigital data that treats data transformation not simply as making ‘end-products,’ but rather as a process that generates intermediate data sets, i.e. datafication, within the dynamics of archaeological practice. In this way, as we transform and integrate analog and digital data, we acquire new knowledge about data collection, documentation, processing, and interpretation that can lead to new archaeological questions and methodologies and enhance the nature of our scholarship.

Data Accessibility Statements

The geospatial data are available to researchers upon request via the MayaArch3D and MayaCityBuilder Projects, and with approval from the Instituto Hondureño de Antropología e Historia (IHAH).

Acknowledgements

We would like to thank ARKWORK COST Action WG4 on ‘Archaeological Scholarship’ for funding the inspirational workshop in Cologne, Germany. We also grateful to Special Issue editors Eleftheria Paliou, Jeremy Huggett, Costas Papadopoulos, and Isto Huvela, as well as workshop participants, the Honduran Institute for Anthropology and History (IHAH), the MayaArch3D Project, and the anonymous reviewers.

This article is based upon work from COST Action ARKWORK, supported by COST (European Cooperation in Science and Technology), www.cost.eu.

Competition Interests

The authors have no competing interests to declare.

References


Alberts, G, Went, M and Jansma, R. 2017. Archaeology of the Amsterdam Digital City; Why Digital Data are


Beale, G and Reilly, P. 2017. Digital Practice as Meaning Making in Archaeology. Internet Archaeology, 44. DOI: https://doi.org/10.11141/ia.44.13


Gillings, M and Goodrick, GT. 1996. Sensuous and Reflexive GIS: Exploring Visualisation and VRML. Internet Archaeology, 1. DOI: https://doi.org/10.11141/ia.1.2


Huggett, J. 2017. The Apparatus of Digital Archaeology. Internet Archaeology, 44. DOI: https://doi.org/10.11141/ia.44.7


Richards-Rissetto and Landau: Digitally-Mediated Practices of Geospatial Archaeological Data


Richards-Rissetto and Landau: Digitally-Mediated Practices of Geospatial Archaeological Data


Wither, RE. 2008. (Re)surveying Mediterranean rural landscapes: GIS and legacy survey data. Internet Archaeology, 24. DOI: https://doi.org/10.11141/ia.24.2


Submitted: 08 January 2019 Accepted: 03 July 2019 Published: 22 August 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Journal of Computer Applications in Archaeology, is a peer-reviewed open access journal published by Ubiquity Press.