December 2022

Virtual Excavations: Digital Repositories, Data Reuse, and Ethically Accessible Archaeology

Allison Lindsey Densmore
University of Wisconsin-Milwaukee

VIRTUAL EXCAVATIONS: DIGITAL REPOSITORIES, DATA REUSE, AND ETHICALLY ACCESSIBLE ARCHEOLOGY

by

Allison Lindsey Densmore

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Anthropology

at

The University of Wisconsin-Milwaukee

December 2022

ABSTRACT

VIRTUAL EXCAVATIONS: DIGITAL REPOSITORIES, DATA REUSE, AND ETHICALLY ACCESSIBLE ARCHEOLOGY

by

Allison Lindsey Densmore

The University of Wisconsin-Milwaukee, 2022
Under the Supervision of Professor R. Jason Sherman

Archaeological investigations produce massive amounts of data, yet these data are often sequestered by the original researchers or put behind paywalls that restrict access to academic publications. This inaccessibility makes it difficult to justify the destructive nature of archaeology. Open-access digital data management systems such as the Digital Archaeological Record (tDAR) provide archaeologists with new ways to preserve, share, and, most importantly, reuse archaeological data to combat this issue of data sustainability. The goal of this project is to investigate how these digital repositories allow for ethically responsible data access and reuse, thus mitigating the cycle of destruction, hoarding, and inaccessibility. Two tDAR-based case studies form the core of this research: the Mimbres Pottery Images Digital Database (MimPIDD) and the Salt River Project Digital Library (SRPDL). This study examines the structure, content, accessibility, and instances of reuse of the case studies as well as quantifiable interactions with their content, including view and download patterns. These analytics are discussed alongside the FAIR and CARE Principles of digital stewardship and the guidelines for data access and management established by stakeholders. By analyzing these factors alongside specific instances of data reuse in academic and public spheres, this project demonstrates how MimPIDD and SRPDL use digital dissemination platforms to create opportunities for meaningful interactions

with their data. These accessible and reusable projects should serve as blueprints for the future of ethical and accessible data management in archaeological research and beyond.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. R. Jason Sherman, Associate Professor in the Department of Anthropology at UW-Milwaukee (UWM), who believed in and supported me through every step of developing this thesis. I am also very grateful for my committee members, Dr. Bettina Arnold, Professor in the Department of Anthropology at UWM, who gave me the encouragement needed to kickstart what I saw as a pipe dream of a thesis, and Dr. Kevin Garstki, Lecturer in the Department of Anthropology, Global Religions, and Cultures at UW-Oshkosh, whose course *Digital Archaeology and Data Reuse* opened my eyes to the amazing world of digital archaeology.

I would like to extend my deepest gratitude towards Rachel Fernandez, Digital Preservation Program Manager at Digital Antiquity at Arizona State University (ASU), and Dr. Christopher Nicholson, Director of Digital Antiquity at ASU, for meeting with me, answering countless questions over email, and supplying invaluable data used in this research. I would also like to thank Dan Garcia, Senior Cultural Resource Management Specialist for the Salt River Project, for taking time to meet with me and providing beneficial insights into the Salt River Project Digital Library. Additionally, I would like to thank MimPIDD board members Dr. Michelle Hegmon, Professor in the School of Human Evolution and Social Change at ASU, and Dr. Will Russell, current head of the MimPIDD board and Historic Preservation Specialist for the Arizona Department of Transportation, for sharing key data on the reuse of the Mimbres Pottery Images Digital Database.

Finally, I would like to thank my family and friends for their endless support. I would not have been able to complete this thesis if it weren't for the love and encouragement of my incredible family and friends, who were always willing to listen to me talk endlessly about the

ever-exciting topic of digital data management. I am also extremely grateful for my cohort of graduate students in the Anthropology Department at UWM for their feedback and inspiration since the inception of my research topic and sense of community that assured me that I was not alone throughout this process. Thank you.

Chapter 1: Introduction

This thesis focuses on two case studies, the Mimbres Pottery Images Digital Database (MimPIDD) (tDAR id:22070) and the Salt River Project (SRP) Digital Library (tDAR id:57630), to illustrate how digital archaeological data repositories make ethically responsible data access and reuse possible. This study examines the structure, content, accessibility, and instances of reuse of each case study to demonstrate how both MimPIDD and the SRP Digital Library use digital data dissemination platforms to create opportunities for meaningful interactions with their data.

MimPIDD and the SRP Digital Library are collections housed in tDAR (the Digital Archaeological Record), a digital repository for archaeological data. MimPIDD contains nearly 10,000 resources, mostly images, of pottery vessels from the Mimbres cultural region in the American Southwest. The database has both public and research versions; the former is accessible to anyone, whereas the latter requires a written request to access. The two versions differ in the number of images and attributes associated with each vessel, primarily to protect sensitive spatial and burial data. MimPIDD emphasizes the goals of accessibility through its curated K-through-college educational activities.

The SRP Digital Library differs from MimPIDD in that it contains gray literature and technical reports from across Arizona. The resources within the SRP Digital Library are primarily used by cultural resource management (CRM) professionals. There are over 700 files in the SRP Digital Library that users can request access to on a file-by-file basis. Some resources within the SRP Digital Library have additional redacted versions of the same files that protect specific locational information and culturally sensitive imagery. These redacted files do not require permission to access. Requests to access the confidential resources within the SRP

Digital Library are granted based on the legitimacy of the research proposal as outlined by the user in their request. This ensures that only those who need the confidential data are given access to them and that any sensitive information does not get into the wrong hands. For example, if a user requests access for personal use or for research that does not necessitate use of the unredacted portions of the document, their request will be rejected. As discussed in Chapter 2, decisions surrounding confidential data should be made in collaboration with stakeholders and those entrusted with stewardship of the data.

Data Reuse in Archaeology

Archaeological materials, whether in the form of physical artifacts or tabulated data, are non-renewable resources. As soon as the first trowel hits the soil, the archaeological materials and surrounding matrix are disturbed and can never return to their original state. The inherently destructive nature of archaeological excavations, as well as the massive outputs of data that result from this work, create a complex problem for archaeological data management. This problem is further exacerbated by what Cherry (2011:17) calls a "crisis of confidence" that motivates archaeologists to keep digging and recovering more material even though vast amounts of data are already available. To justify this destruction of the archaeological record and the massive scale of existing legacy data, archaeologists must make the most of these data through access and reuse (Fagan 1995:16). Digital data repositories provide a means for doing so.

Much of the published literature in archaeology is written for academic audiences, limiting who can interact with it in the first place. Many excavations and other research projects are never published at all, and raw data and preliminary reports are often sequestered by

researchers. Some of this is due to the "publish or perish" mindset in academia, which limits sharing work with others due to fears of being out-published (Beck and Neylon 2012:486; Kansa and Kansa 2013:88; Richards 2015:66), but it is also the case that good field archaeologists are not always gifted writers, and vice versa. This longstanding linear-flow model of data publication (Figure 1.1) is not necessarily conducive to a collaborative and resource-conscious approach to archaeology. One of the goals of this thesis is to suggest how this situation can be improved by adopting a sustainable approach to data publication that takes full advantage of the digital data management tools that are now available to archaeologists (Figure 1.2).

Compared to the typical dead-end model of data publication, the improved approach is both accessible and cyclical. The raw data originate in the same way as before —with the recovery of physical remains, the creation of the data, and the subsequent destruction of context —but the flow of information changes directions by being published in accessible digital data repositories. At this point in the publication process the nature of access is fundamentally different from that of the previous model. Reused data can then be published and put back into an accessible repository, fulfilling the cycle and creating more opportunities for reuse and collaborative contributions. Moreover, data reuse is not restricted to repurposing legacy datasets, but can also be seen in teaching, public-facing blogs, museum exhibits, and heritage signage.

Figure 1.1. Typical model of archaeological data publication.

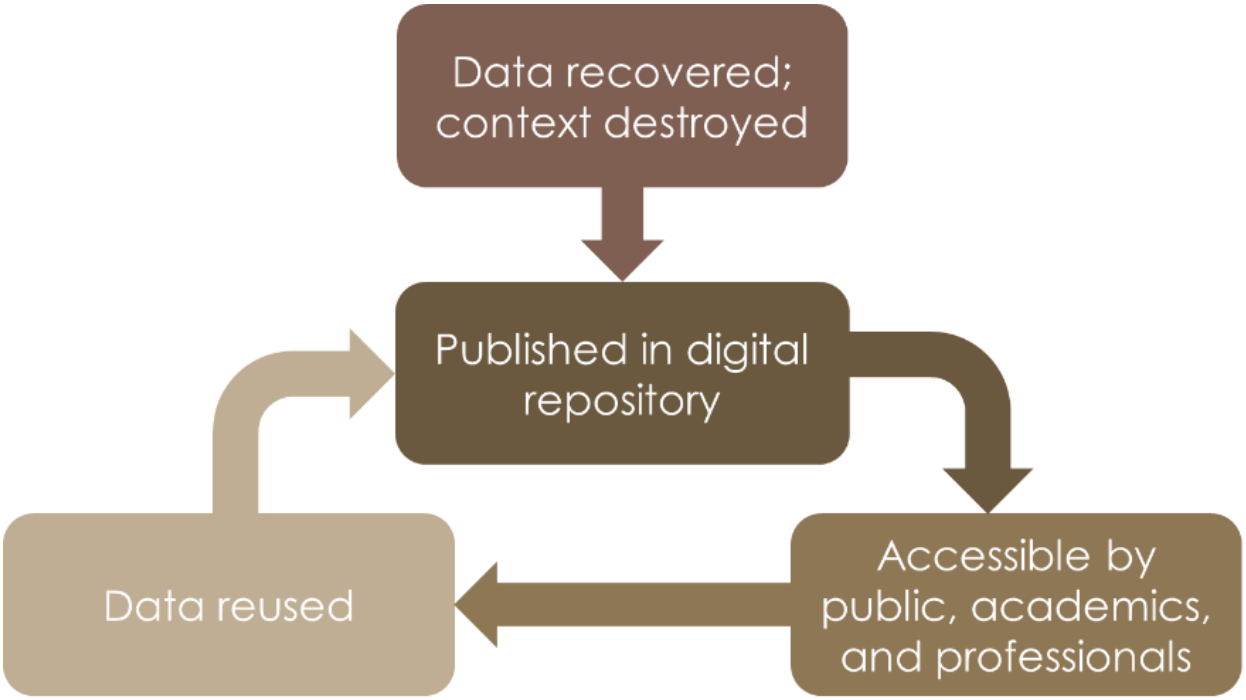

Figure 1.2. Improved model for archaeological data publication.

Those involved in digital data management today include repositories like tDAR and the Archaeological Data Service (ADS), data publishers such as Open Context, and ARIADNEplus, a framework that connects data across repositories. These data repositories, publishers, and frameworks can house virtually any file type including documents, datasets, images, and geospatial files, meaning practically all data can be made accessible on these platforms. Data accessibility is the cornerstone of these repositories. Their focus on accessibility allows for public, academic, and professional engagement with the materials. This, in turn, encourages and facilitates data reuse, so the cycle of access and reuse can continue to repeat itself (Huggett 2018:96).

Archaeological data reuse ensures that archaeologists go beyond simply breaking even in making destructive excavations and time-intensive research "worth it" by only publishing preliminary reports without definitive results before moving on to their next research projects (Fagan 1995:15). Thus, as forecast by Fagan (1995), digital data publication and the resulting capacity for data reuse allows archaeologists to not only slow down and focus on interpreting already existing data, but also to continue to make significant contributions to science through multiple perspectives and by making connections between otherwise isolated datasets (Currie 2014:148; Fagan 1995:17-18; Wylie 2017:8).

The "publish or perish" mindset mentioned above is engrained in the very fabric of academia. Therefore, a shift towards open data could raise concerns for academic researchers and their goal of sustainable relevancy. However, with appropriate documentation and metadata, academic researchers can share their data without feeding into the fear of perishing. Digital Object Identifiers (DOIs), ORCID identifiers, and other permanent identifiers help to ensure that credit is maintained and that researchers will still be tied to their original data, partially ameliorating the publish or perish dilemma. These persistent identifiers work by maintaining unique codes to refer to specific publications and researchers, guaranteeing their findability (Koster and Woutersen-Windhouwer 2018). Equally weighing the value of data sharing alongside typical academic publications could also promote a beneficial shift in the publish or perish environment, allowing researchers to focus on what truly matters: the data itself.

The impact of digital data in archaeology is greatly affected by ethical constraints. The manager, gatekeeper, and custodian roles that archaeologists must play when dealing with sensitive cultural information become even more important when data are made public (Richardson 2018:64). Kansa (2016:459) writes, "treating data as yet another research product

needing to be managed and measured undermines both intellectual freedom and the ethical conduct of research." How then can and should archaeologists ethically manage their data? It is important to consider that regulations about how to protect and ensure the viability of data vary globally, so archaeologists must be conscious of the relevant guidelines applicable to their particular project area.

Data do not become reusable just because they are uploaded to digital repositories. For data to be not only accessible but also reusable by others, they must adhere to certain criteria. The FAIR Guiding Principles for scientific data management and stewardship—including, **F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability—provide standards for data publication to improve the capacity for data reuse and sustainability (Wilkinson et al. 2016:1). If these standards are not met, the capacity for high-yielding reuse is greatly diminished. The FAIR Principles are complemented by the CARE Principles for Indigenous Data Governance which emphasize **C**ollective benefit, **A**uthority to control, **R**esponsibility, and **E**thics (Carroll et al. 2020:1). While the CARE Principles were created in the context of Indigenous data management, they can and should be extended to all archaeological data management discussions. Together, the FAIR and CARE Principles form the theoretical basis of this project as they represent the ideal qualities of reusable digital archaeological data. The FAIR and CARE Principles will be discussed in greater depth in Chapter 2.

The advent of digital data management tools in archaeology promises to unlock the incredible potential of the discipline for understanding human culture through space and time. This change is so significant that it has been considered a paradigm shift by many (e.g., Huggett 2020; Roosevelt et al. 2015:339). With born-digital technologies becoming more accessible and the wave of digitization sweeping the discipline, it is likely that the discipline will be fully digital

in the next few generations as training in digital data management increases in tandem with traditional academic research.

Research Goals and Questions

This thesis project originated with the recognition that while digital data repositories are changing the way archaeological data are published, accessed, and reused, these platforms are relatively new to archaeology. This means that there have been few investigations into the quantity, quality, and varieties of archaeological data reuse, specifically data sourced from digital repositories. Moreover, there is no standard methodology for analyzing and interpreting reuse and based on conversations with the directors of Open Context and tDAR, this is something that many of the major repositories are trying to develop. This thesis aims to fill this methodological gap.

In this thesis I investigate the structure, content, and accessibility of two collections in tDAR, and how these characteristics influence how data are being reused (if at all) by evaluating specific instances of reuse. The research questions I address include:

- As self-appointed stewards of the archaeological record, how can archaeologists use digital data management systems to reconcile public reporting requirements with the need to ensure the protection of sensitive cultural materials?
- The FAIR and CARE Principles outline the ideals for data management, with the CARE Principles, in particular, focusing on ethical data management. How can archaeologists ensure that the FAIR and CARE Principles are being followed when dealing with digital archaeological data?

- If the databases contain sensitive data, what are the best practices for managing access to the materials?
- How do the structure, content, and accessibility of digital archaeological databases affect the reusability of data?
- What can view and download metric files suggest about how dual-access databases are accessed and reused by professionals and the public?
- How can archaeologists quantitatively and qualitatively measure the reuse of archaeological data?

The question of how to measure archaeological data reuse is the crux of this thesis. While numbers of views and downloads can suggest how data are being accessed in digital repositories, they do not necessarily indicate whether and how data are being reused. As Huggett (2018:94) suggests, "such metrics primarily capture supply rather than reuse." Qualitative reuse can be difficult to measure as it occurs beyond the bounds of the repositories themselves. Reuse is thus hard to track aside from citations in academic publications, or direct feedback from users in the form of direct personal communication or online surveys, as discussed in Geser et al. (2022). Nonetheless, reuse metrics can serve as an important tool for gauging the shift towards this cyclical and accessible data framework, so it is beneficial to develop viable methods for gathering data on archaeological data reuse.

Thesis Significance

This thesis contributes to the growing discussion about ethical and practical considerations involved in digital data publication and reuse in archaeology. Additionally, the conclusions add to the broader discussion about data management across other ethically- and

resource-conscious disciplines. Digital data publication and the resulting prospect of reuse are topics that few archaeologists are aware of. I argue that this is something that needs to be at the forefront of archaeologists' minds before, during, and after the data collection process.

I hope that the methods used in this thesis to measure the accessibility and instances of reuse of the case studies can be applied in similar studies or internal investigations within archaeological databases and/or repositories. The number of researchers focusing on digital archaeological data access and reuse continues to expand and so do the research questions associated with this phenomenon. This thesis is intended to be a steppingstone towards developing an accessible digital archaeology. The research presented in this thesis can also be of use to the administrators of MimPIDD and the SRP Digital Library, as well as tDAR more generally, by providing summative assessments on the accessibility and reusability of their data. I hope this thesis demonstrates that the work that went into creating and managing these and similar projects is worthwhile and has tangible benefits for the creators and other potential stakeholders.

Thesis Outline

Chapter 2 of this thesis provides background information on the transition from analog to digital data management in archaeology. This chapter contains a detailed explanation of the FAIR and CARE Principles introduced in Chapter 1. It continues with an overview of digital data management systems in archaeology and the importance of metadata in ensuring accessibility and reusability of data. The chapter concludes with an in-depth introduction to the case studies, MimPIDD and the SRP Digital Library.

Chapter 3 describes the methodology and data used in this thesis. It provides an overview of the methods with subsections that describe how the structure, content, accessibility, and reusability of MimPIDD, the SRP Digital Library, and tDAR as a whole are measured. This chapter also explains how the data on each case study was sourced.

Chapter 4 presents the results and analysis of the quantitative analyses of views and downloads for a sample of resources within both MimPIDD and the SRP Digital Library. This chapter also includes qualitative analyses of reuse of the two databases. Chapter 4 ends with a section on overall trends in tDAR.

The conclusions resulting from the analyses presented in Chapter 4 are outlined in Chapter 5, which comprises five sections. The first section provides a summary of the analyses conducted for this thesis. The second and third sections address the research questions posed in Chapter 1 with suggestions for how archaeologists can better transition into a digital archaeology. The fourth section discusses the significance of the research presented in this thesis. Finally, the chapter poses future research directions regarding digital data access and reuse in archaeology.

Chapter 2: Background

The Curation Crisis and the Digital Dilemma

As far back as the 1980s (e.g., Marquardt et al. 1982), concern for the state of archaeological collections has been at the forefront of archaeological discussions. This concern has only become exponentially more important as the discipline shifts towards the digital realm. Kersel (2015) and others have been actively debating how we can approach this growing divide between the amount of data collected and the time, space, and funding available to deal with this information overload. She makes an important point in saying, "The very term *curation* implies careful stewardship of collections; management that includes accessioning, cataloging, conserving, maintaining, processing, publishing, and storing artifacts and the associated documentation" (Kersel 2015:42).

While the curation crisis began as an analog issue well before the large-scale integration of digital technologies, it is something that has inevitably carried over into the digital age. Digitizing the curation process as well as archaeological materials and data themselves can have significant benefits in terms of querying, analysis, and organization in general, making it the obvious next step in collections management (Wallrodt 2016:46). Clarke (2015:314) characterizes this new era as the "Digital Dilemma," specifically bringing attention to the new pressures presented by the fact that "if scholars wish to ensure the long-term viability of their digital data, they must navigate the potential conflicts that arise from these advances."

In recent years, there has been a noticeable shift towards a digital archaeology as photogrammetry, handheld internet-based GPS technologies, and digital field data collection forms have taken over as mainstream data collection techniques. These are all examples of born-

digital data, or data collected in digital rather than analog form (Garstki 2020:6). As Roosevelt et al. (2015:339) write, "this shift mean(s) avoidance not only of paper, but also of tape measures, line-levels, and drafting sheets." By taking advantage of these new applications of software we can produce a "paperless workflow" of data which allows for an increased pace in both the collection of and access to archaeological information (Wallrodt 2016: 38). The shift toward more digital technology in archaeology does not necessarily require changing every aspect of archaeological research, but at the minimum, archaeologists should be aware of the potential benefits of digitization and born-digital data in the context of their research projects. A simple way to integrate a digital data is through digitizing collections catalogs, making them more findable than their analog or otherwise inaccessible counterparts. These could allow archaeologists to track down otherwise hard-to-find resources in their own collections and beyond.

Digital data do not have to be born-digital but can include the digitized version of analog notes, maps, inventories, and more. These types of digital data have the same, if not greater, potential to make a difference in the practice of archaeology. Born-digital data and digitization are changing the way that archaeologists amass data. This means that the data can get overwhelmingly large rather quickly. Take, for example, the increasing use of high-resolution aerial and satellite imagery in archaeological investigations. These images and other Big Data, characterized by their scale and general inability to be processed by humans alone (VanValkenburgh and Dufton 2020:S2), are a part of what Bevan (2015:1480) refers to as the "data deluge." The "chronic annual flooding" of datasets, raw images, gray literature, and more necessitate an urgent shift in digital archaeological data management (Bevan 2015:1480). It is

this very accumulation of information that has accelerated the transition to a fully digital archaeology.

Open Archeology

Open data are characterized by their accessibility, technical openness, and legal openness (Kansa and Kansa 2013:94). Open data are invaluable in the current age of research as they allow for collaboration in a way that was not possible before. However, simply because data are open does not mean they are automatically ready to be reused. In the new world of open data, Big Data, and mass digitization, it remains important to consider the universal qualities of data. While the term "raw data" often gets thrown around when referring to these Big Data, it is vital that we acknowledge the fact that data are inherently shadowy and "theory-laden, process-laden, and purpose-laden, and not raw in any sense" (Huggett 2020:S9; also, Shanks and Tilley 1992:37; Wylie 2017). To keep track of these factors of scaffolding, archaeologists must keep accurate and detailed accounts of their data collection methods, a "biography of data" (Huggett 2020:S14; also, Roosevelt et al. 2015:325). Massive amounts of data are being created and thus need to be stored and maintained with their biographies intact. Big Data and the data deluge itself are inherently messy, making that a difficult job for archaeologists (Gattiglia 2015:114). Digital data repositories and publication services use permanent object identifiers and standardized coding schemes to provide a means of handling these large data.

Due to the breadth of archaeological data, it is not surprising that universal standardization in terms of recordation and publication is incredibly difficult to achieve. To further ensure openness and accommodate the dynamic nature of archaeological data these workflows should be interoperable and accompanied by detailed metadata, ontologies, and

coding sheets (Beck and Neylon 2012:482; Faniel et al. 2018:114, Isaksen et al. 2009:7). Metadata are descriptions of and information about data such as the contributors, file format, date of creation, and operating system(s) used to create a file (Clarke 2015:317-318; Huggett 2020:S12). Ontologies are data models that hierarchically organize data values and allow users to link datasets and coding sheets (Digital Antiquity 2020). These behind-the-scenes aspects of digital data management become more important as increasing numbers of people gain access to the data down the line. If metadata are properly recorded alongside their digital data counterparts, the capacity for data reuse increases exponentially.

Beyond simply having accessible repositories with consistent data, how can we ensure they are used to the fullest extent possible? Wylie (2017) spells out three ways archaeologists can put old data to work through secondary retrieval, recontextualizing data, and experimental simulation. What remains clear is that data never really run dry. A singular database can be used in an infinite number of ways to answer a great variety of questions. When you can combine one database with another, the possibilities multiply exponentially. In any case, the output of data reuse is only as good as the input, especially when combining multiple datasets. Thus, high levels of standardization and organization are necessary for representational and meaningful results (Clarke 2015:315). There is, however, an extent to which standardization could do more harm than good. Thus, standardization must be structured enough to allow for interoperability without being so restrictive that it limits potential conclusions. Regardless, a "slow data" approach wherein researchers emphasize the process of creating data and the implications that follow can ensure that data are used not just as many times as possible, but with as much depth and consideration as possible (Huggett 2022:98, Kansa 2016:466).

While the notion of having all these data available to other researchers is encouraging, without appropriate intervention, archaeology could be heading towards a “digital Dark Age” wherein data obsolescence and the subsequent loss of unreproducible data becomes unavoidable (Richards et al. 2021). Data need to be properly organized and curated alongside detailed metadata to ensure their survival (Clarke 2015: 315). If archaeologists are unable or unwilling to put sufficient time and resources into taking care of their data, they are actively contributing to the destruction of cultural heritage. The issue of how to manage data and deal with related ethical issues is complex, and if not addressed appropriately, has the potential to create many roadblocks as the field moves forward into a digital era.

The Ethics of Digital Data Management

Archaeology is not immune to ethical dilemmas. The handling of cultural materials, physical or digital, necessitates great care and cultural awareness. Thus, ethical considerations should be at the forefront of every archaeologist’s mind at every stage of data creation, management, and publication (Richardson 2018:65). One of the biggest hurdles in the race towards a digital archaeology is understanding and addressing the ethical issues inherent in this new age of research. As emphasized by Dennis (2020:210), there exist clear expectations for conventional (analog) archaeological ethics but vague guides for ethical decision making in the digital realm where the ethical waters are already muddied.

Stewardship is undeniably important in every archaeological context but is foundational to digital data management. According to the Society for American Archaeology (SAA), whether dealing with people, sites, or artifacts, archaeologists must ensure that they are "caretakers of and advocates for the archaeological record" (Society for American Archaeology 2018). This same

emphasis on stewardship and dissemination is reiterated by the Archaeological Institute of America (AIA) in their *Code of Professional Standards*. Both organizations place stewardship and public involvement at the forefront of their ethical standards. The ethicality of archaeology and data management are contingent on this balance between protecting and publicizing archaeological data. Nevertheless, there is an undeniable "stewardship gap" in archaeology (York et al. 2016).

Many archaeological investigations are publicly funded, conducted on federal land, or concern the history of the collective human past, with over 90 percent of excavations in the US being done pursuant to Section 106 of the National Historic Preservation Act (NHPA) by cultural resource management (CRM) firms (Advisory Council on Historic Preservation 2018). It follows that the data recovered should be made as accessible as possible (Beck and Neylon 2013:485). As Altschul and Klein (2022:13) suggest, CRM documents the past but leaves it largely uninterpreted. Amplified by a lack of data systematization and the general inaccessibility of CRM gray literature and associated data, the vast backlog of existing data recovered by CRM archaeology fails to meet all four of the FAIR Principles (Altschul and Klein 2022:13). CRM funding is predicted to increase from $1.46 to $1.85 billion between US fiscal years 2021 and 2031 (Altshcul and Klein 2022:1). This increase in funding and growth of the field provides an opportunity to allocate funds to resolve the CRM-related digital data management issue.

While digital data management provides a means for making data accessible, there are ethical and financial considerations that must be understood before data are uploaded to a digital data infrastructure. Dennis (2020:215) argues that not only should ethical guidelines for digital curation be considered, but also how those digital data relate to living people. Nicholson et al. (2021) and Richardson (2018:70) recommend various ways that archaeologists can improve the

preservation of and access to digital archaeological data. Their recommendations include incentivizing archiving practices in academia by educating faculty and institutions on its advantages, promoting data preservation standards and including data curation in the curriculum for undergraduate and graduate archaeology students, and having professional archaeological organizations provide archaeologists with ethical and practical guidelines related to data archiving (Nicholson et al. 2021). Currently there is insufficient funding for post-excavation work, let alone digital storage and maintenance. For the aforementioned recommendations to be implemented, funding focused specifically on digital data management must be increased. Along with the FAIR and CARE Principles discussed below, these recommendations can serve as a framework to ensure the viability and ethicality of data sharing.

Accessible Public Archaeology

As mentioned above, archaeologists should look beyond giving access to fellow academics to further archaeological research; they should also work towards an open and public archaeology. Public archaeology can include, "the democratization of communication, activity or administration; communication with the public; involvement of the public, or preservation and administration of archaeological resources for the public benefit by voluntary or statutory organizations" (Richardson 2013:2). Accessible and public-facing archaeology focuses on the goals of direct involvement with archaeological materials and practices as well as the general sharing of information.

Getting the public involved in and aware of archaeological information helps connect them to the collective human past and thus come to respect and care for archaeological work. Bonacchi and Moshenshka (2015:2) identify seven common types of public archaeology:

archaeologists working with the public, archaeology by the public, public sector archaeology, archaeological education, open archaeology, popular archaeology, and academic public archaeology. Digital public archaeology can benefit the public and professionals alike through mutually beneficial interest and engagement (Bollwerk 2015:228; Richardson 2018:70).

Opportunities for public engagement with archaeology can vary from hands-on excavation experiences through programs like DigVentures to data sharing through databases like the Portable Antiquities Scheme (PAS). DigVentures is a crowdfunding and crowdsourcing program that allows the public to participate in legitimate archaeological excavations, covering four of the nine types of public archaeology outlined by Bonacchi and Moshenshka (2015). The PAS, managed by the British Museum, similarly works to increase public engagement with archaeological materials by encouraging the public to share their metal-detecting finds in an online database. As of October 2, 2022, the PAS website boasted that it contained information about more than 1.6 million objects. These hands-on involvement projects go together with Indigenous-focused public engagement programs and educational opportunities like museum exhibits and digital data repositories to educate the public while also providing archaeologists with new and valuable perspectives (Bria and Vasquez 2022:61; Perry 2018:220).

It is important to consider exactly how data will be presented when sharing them with the public. Digital public archaeology has the capacity to reach many more people than a museum, for example, making it even more critical to ensure that information is protected properly. There are inevitably some data that are more sensitive than others and it is important to keep this in mind when publishing or otherwise disseminating information. This necessitates prioritizing important ethical considerations to protect culturally sensitive data.

The FAIR Principles

The FAIR Guiding Principles for Scientific Data Management and Stewardship were first drafted at a workshop in 2014 by a group of stakeholders from various disciplines with the goal of making data, and digital scholarly objects more broadly, easily findable and reusable by humans and machines (Wilkinson et al. 2016:3). As mentioned in Chapter 1, the FAIR Principles include Findability, Accessibility, Interoperability, and Reusability. These four characteristics, outlined in detail in Figure 2.1, are key to maximizing the potential of digital data management.

When followed, the FAIR Principles can provide researchers with new ways of interacting with digital data due to their machine-actionability. Computers can reference, retrieve, and analyze data in ways that humans alone cannot (Boeckhout et al. 2018:933). From a FAIR perspective, physical collections are often "…inaccessible and unusable backlog(s)" that could face obsolescence if they are not discoverable in the first place (McManamon et al. 2017:240). Housing reports, images, maps, and other data in digital data management systems alongside detailed metadata enhances the research potential of physical collections by ensuring that they are findable, accessible, interoperable, and reusable. There is no question that physical collections are vital to archaeological research, but they cannot serve a role in a reuse-focused archaeology if researchers do not know that they exist. The findability, or lack thereof, of an analog collection is itself disadvantageous, making the AIR of FAIR even harder to achieve. In this way, the FAIR Principles build upon each other; accessibility is predicated on findability, interoperability is predicated on accessibility, and reusability is predicated on interoperability.

The FAIR Principles are intentionally broad to make them applicable to different data management systems across disciplines with varying types of data and necessary levels of metadata (Wilkinson et al. 2016:4). While this is a great strength of the FAIR Principles, problems could arise without discipline-specific supplemental principles (Boeckhout et al. 2018:935). There are certain data and methods specific to archaeology that can complicate adherence to the FAIR Principles (Richards et al. 2021). Archaeology deals with a wide range of data types, many of which necessitate greater levels of protection due to their culturally sensitive content. This is where the CARE Principles, discussed below, and the notion of keeping data as open as possible and as closed as necessary, become vital.

Findable

- **F1**. (Meta)data are assigned a globally unique and persistent identifier
- **F2**. Data are described with rich metadata (defined by R1 below)
- **F3**. Metadata clearly and explicitly include the identifier of the data it describes
- **F4.** (Meta)data are registered or indexed in a searchable resource

Accessible

- **A1.** (Meta)data are retrievable by their identifier using a standardized communications protocol
- **A1.1.** The protocol is open, free, and universally implementable
- **A1.2.** The protocol allows for an authentication and authorization procedure, where necessary
- **A2.** Metadata are accessible, even when the data are no longer available

Interoperable

- **I1.** (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- **I2.** (Meta)data use vocabularies that follow FAIR principles
- **I3.** (Meta)data include qualified references to other (meta)data

Reusable

- **R1.** Meta(data) are richly described with a plurality of accurate and relevant attributes
- **R1.1.** (Meta)data are released with a clear and accessible data usage license
- **R1.2.** (Meta)data are associated with detailed provenance
- **R1.3.** (Meta)data meet domain-relevant community standards

Figure 2.1. The FAIR Guiding Principles for scientific data management and stewardship (adapted from Wilkinson et al. 2016: Box 2).

The CARE Principles

The CARE Principles for Indigenous Data Governance were developed by the International Indigenous Data Sovereignty Interest Group within the Research Data Alliance as a people- and purpose-oriented complement to the FAIR Principles (Carroll et al. 2020:3). The CARE Principles (outlined in detail in Figure 2.2) include four key characteristics: Collective benefit, Authority to control, Responsibility, and Ethics. These four principles "define rights, interests, and concepts to be employed in facilitating control in data governance and reuse" (Carroll et al. 2020:5).

Colonization, and the subsequent exploitation of Indigenous data, are practically foundational to archaeology in the United States and beyond (Nicholas and Hollowell 2007:60; Shanks and Tilley 1992:28; Trigger 1984:361). According to Carroll et al. (2020:3), Indigenous data comprise:

> (1) information and knowledge about the environment, lands, skies, resources, and non-humans with which they have relations; (2) information about Indigenous persons such as administrative, census, health, social, commercial, and corporate and, (3) information and knowledge about Indigenous Peoples as collectives, including traditional and cultural information, oral histories, ancestral and clan knowledge, cultural sites, and stories, belongings.

Thus, the ethics-focused CARE Principles are especially relevant to aspects of archaeological data management where the FAIR Principles fall short. In a follow-up discussion regarding how the FAIR and CARE Principles work together, Carroll et al. (2021:2) assert that they ensure that "… the use of data aligns with Indigenous rights, is as open as determined by Indigenous communities, is purposeful, and enhances the well-being of Indigenous Peoples." As

conversations regarding Indigenous data rights and stewardship continue to be had in archaeology and beyond, "…we must be willing to relinquish control over their cultural heritage" (Supernant and Warrick 2014:581).

Given that the CARE Principles were only established in 2020, they are still being introduced and integrated into archaeology and beyond. As discussed in greater depth in Chapter 4, the two case studies in this thesis both revolve around indigenous data, with MimPIDD focusing on Mimbres pottery and the SRP Digital Library focusing on Hohokam archaeology. Both databases feature differentiated access models to protect culturally sensitive data.

Collective Benefit

- **C1.** For inclusive development and innovation
- **C2**. For improved governance and citizen engagement
- **C3**. For equitable outcomes

Authority to Control

- **A1.** Recognizing rights and interests
- **A2**. Data for governance
- **A3**. Governance of data

Responsibility

- **R1.** For positive relationships
- **R2**. For expanding capability and capacity
- **R3**. For Indigenous languages and worldviews

Ethics

- **E1.** For minimizing harm and maximizing benefit
- **E2.** For justice
- **E3.** For future use

Figure 2.2. The CARE Principles for Indigenous Data Governance (adapted from Carroll et al. 2020: Fig. 2).

Digital Data Management Systems

Beyond ethics, archaeological data are by nature difficult to manage and disseminate given the massive amount of data typically collected from even a single excavation or survey project. Digital data management systems are relatively new to archaeology, and they are still gaining traction and growing in both size and capability. Moreover, many archaeologists are

unaware of the true potential of practicing a digital archaeology, let alone how to work towards it (Faniel et al. 2018:106). Digital data management systems provide a way to share, protect, and preserve these data in responsible ways.

As mentioned in Chapter 1, archaeology-specific digital data repositories and publishers are becoming more mainstream as the digital shift becomes imminent. Repositories like the Archaeological Data Service (ADS) and the Digital Archaeological Record (tDAR), discussed in greater depth in the following section, focus on long-term curation and preservation while Open Context uses a peer-review data publication model. These platforms allow users to upload a wide range of file types, including spreadsheets, GIS files, PDFs, and images. Making raw data available in this way is essential for data access and reuse as it allows academics and the public to engage with otherwise inaccessible information (Atici et al. 2012:670; Bevan 2015:1477-1478).

There is a definite and ever-expanding gap between archaeological education and the actual practice of archaeolgy, especially when it comes to digital data management (Garstki 2022:2). The lack of concern for digital data management in archaeology, as suggested by Nicholson et al. (2021), comes from both the fact that training of young archaeologists emphasizes field methods rather than follow-up activities like report writing and artifact management as well as the lack of enforcemnent of laws and regulations put in place to protect archaeological data. The NRHP and other federal laws such as the Archaeological Resources Protection Act (ARPA) set specific curation requirements for archaeological data, yet they often go unfollowed. 54 USC § 306131(a)(1)(c) of the NRHP dictates that data must be "permanently maintained in appropriate data bases and made available to potential users" and 16 USC § 470cc (b)(3) of ARPA states that data recovered on public land must be "preserved by a suitable

university, museum, or other scientific or educational institution" The NRHP and ARPA serve as the foundation for CRM work in the United States, yet these two requirements are often overlooked. Industry-specific digital data management systems provide a solution to this problem.

Open Context advertises a "data sharing as publication" model (Kansa and Kansa 2013:89). The use of DOIs, globalized ontologies, and peer-review make publications in Open Context comparable to typical academic publications. While the platforms make it easy to preserve and disseminate all types of data, they can cost a considerable amount of money depending on the size of the project, making sharing data in this way difficult for researchers without access to the necessary funds (Kansa 2016:458). Kansa (2016:447) suggests that the complications of this digital, open publication are similar to desperate-for-money low-budget commercial set-ups, differing greatly from other forms of academic publication. Hopefully, a shift from the traditional academic publication model in which only complete reports are published (Figure 1.1) towards a more inclusive model that emphasizes data sharing (Figure 1.2) will be accompanied by an increase in funding opportunities to better support such publications.

Piwowar and Vision (2013:22) report a "statistically well-supported citation benefit from open data"—that is, research using open datasets is cited more frequently than without. Moreover, many repositories track data like the number of views and downloads for a particular resource, shedding light on how it is being accessed. It is important to note, however, that access metrics do not necessarily indicate reuse and a lack of approved methods for quantitatively and qualitatively tracking reuse remains a major roadblock in the quest for a more reuse-focused archaeology (Huggett 2018). Such metrics could be used to replace the traditional publish or perish model.

While funding in archaeology, and the social sciences and humanities more broadly, is already nowhere near ideal, the problem is even more acute with regard to digital data. Time-consuming and sometimes costly work must be done to create databases before they can be uploaded into digital data management systems, which have their own fees. Additionally, digital data have their own curational risks such as obsolete file formats and thus require active curation to ensure reusability (Richards 2002:344). These challeges with curation exist in addition to complex issues regarding who owns data once they are uploaded to a repository, and who is responsible for actively maintaining collections and making decisions about how to best ensure the FAIRness of the data. Nevertheless, these digital dissemination platforms are where archaeologists can balance ethical responsibility and public engagement.

The Digital Archaeological Record

tDAR is a digital repository for archaeological data developed and maintained by the Center for Digital Antiquity (Digital Antiquity), a university center at Arizona State University (ASU). Work on the repository began in the early 2000s with the goal of solving the longstanding issue in archaeology of synthesizing disparate data within a digital infrastructure (Digital Antiquity 2018b). A 2004 workshop funded by the National Science Foundation (NSF) with 31 participants across archaeology and computer science resulted in a report that recommended the development of a "cyberinfrastructure for archaeology" (Kintigh 2006:577). In accordance with these recommendations, the NSF funded the development of such a prototype (McManamon et al. 2021:241). With additional funding from the Andrew W. Mellon Foundation to establish Digital Antiquity, tDAR began to evolve into the repository it is now.

tDAR was designed specifically to accommodate independent data contributions from users around the world. Users can easily upload their resources into tDAR and use built-in resource management tools to organize their records (as discussed in greater detail in Chapter 3). This "self-service digital curation" is enabled by informational resources and instructions such as metadata record templates available on the tDAR website (McManamon et al. 2021:242). Notably, uploading resources to tDAR costs $10 per file (limited to less than 10 MB) or $5 per file if uploading 100 or more files. As of October 6, 2022, the tDAR homepage indicated that the repository contains over 433,000 resources.

As discussed above, DOIs and persistent URLs are vital to a credit-conscious and reuse-focused archaeology. tDAR uses these, in addition to their own tDAR ids, to ensure the FAIRness of the content. tDAR also follows the CARE Principles by not showing the precise location of legally protected resources in the geographic search feature and in the displays on resources pages. Further, and as discussed below and in Chapter 4, tDAR incorporates a user-designated access structure to allow archaeologists to effectively balance both access and protection (Figure 2.3) (McManamon et al. 2017:245).

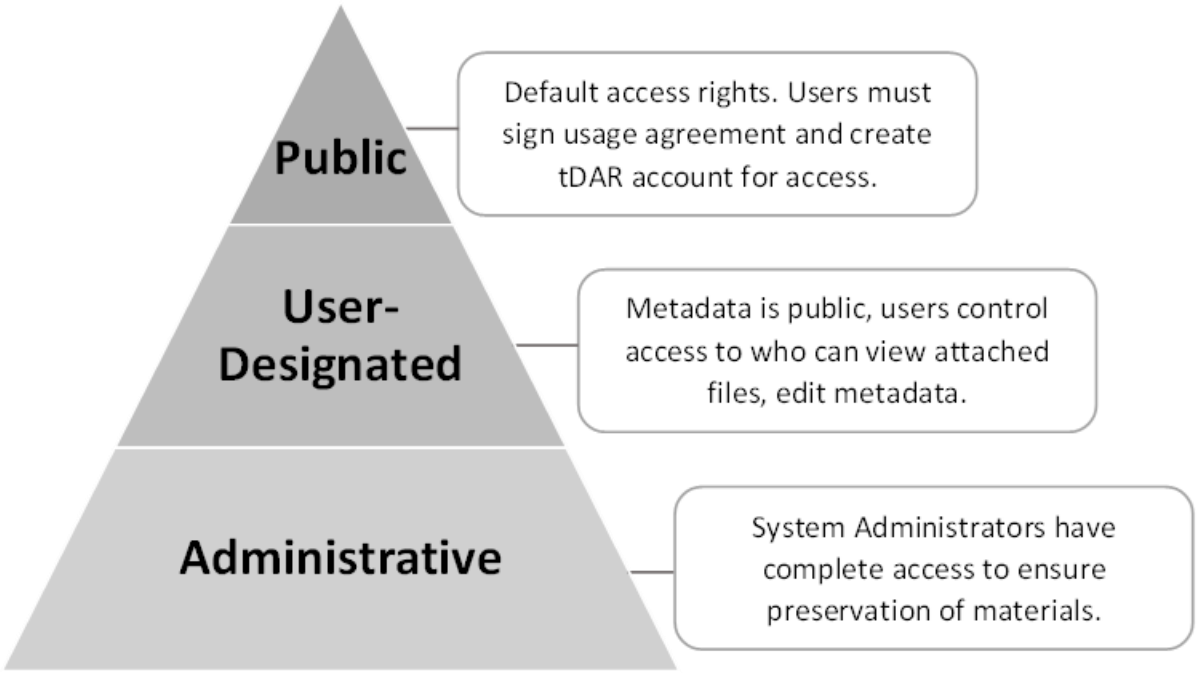

Figure 2.3. tDAR data access framework (adapted from Digital Antiquity 2011).

Introduction to Case Studies

This thesis focuses on two case studies: MimPIDD and the SRP Digital Library. These two databases were chosen because they make different levels of information available to researchers and the public while adhering to the ethical practices discussed above. The following sections discuss the background and development of MimPIDD and the SRP Digital Library. Chapter 4 contains an in-depth analysis of the structure, content, accessibility, and reuse of the two databases.

Mimbres Pottery Images Digital Database

The first database that I analyze in this thesis is the Mimbres Pottery Images Digital Database (MimPIDD) (tDAR id:22070). The database contains attribute data for and images of

over 10,500 Mimbres pottery vessels, mostly Style III bowls. The Mimbres archaeological culture is centered in southwest New Mexico and includes portions of the surrounding American and Mexican states. Mimbres archaeology is best known for its Classic period (100-1130 CE), during which the black-on-white painted pottery was produced that has been a focus of archaeological and art historical interest (Nelson and Gilman 2017:265).

The eye-catching designs on Classic Mimbres pottery include both representational and geometric styles that have garnered interest from archaeologists and the public alike for nearly a century. Many vessels in both the public and the research versions of MimPIDD come from burial contexts or have burial-related characteristics such as kill holes. Kill holes are directly associated with burial ceremonies, but even those vessels without kill holes are likely from burial contexts as they were often placed upside-down on the faces of the deceased during burial ceremonies (Brody 2004:49).

A substantial number of known Mimbres vessels are currently in private collections (Brody 2004:1). There was a dramatic increase in the market value of looted Mimbres vessels when archaeologists began to publish images of the striking bowls in site reports and academic articles, leading many archaeologists to include only drawings of the designs instead (Finegold 2019:231). As a result, nearly every known Mimbres site has been subject to devastating looting and destruction (Brody 2004:10). The high demand for the beautiful vessels was met with attempts by private art dealers to increase the price of their vessels by repainting them, regardless of how accurately done, with more complex and eye-catching designs (Hegmon et al. 2017:7).

The creators of MimPIDD designed the database with the dispersal of artifacts as a result of the art collecting industry in mind. They created the database to assemble as many images and descriptions of Mimbres pottery vessels as possible. A large-scale collection like MimPIDD

would finally allow archaeologists and other stakeholders to perform comprehensive analyses of the now-scattered Mimbres pottery (Hegmon et al. 2017:2). In addition to their goal of facilitating easier large-scale analyses of the vessels, the database's creators are also dedicated to public outreach and education. The online searchable feature allows for users of all backgrounds to explore the database. Users can easily search for various attributes, including archaeological sites, specific designs, and vessel size. Additionally, MimPIDD contains over two thousand images and K-through-college level educational lessons that can be downloaded by anyone with a free and easy-to-setup tDAR login. MimPIDD's content is discussed in greater detail in Chapter 4.

MimPIDD is especially relevant to the goal of balancing ethical obligations and public outreach because there are two levels of access to the database: a public version and a research version. The differences between the two versions are discussed in depth in Chapter 4. Considering how MimPIDD differentiates the two datasets, it is clear that ethical concerns about protecting burial contexts and the risk of encouraging looting and artifact trafficking that comes with private collections warrant an extra layer of protection. After all, stewardship and accountability are vital ethical principles that archaeologists must continue to be aware of even after data collection.

Salt River Project Digital Library

The second case study for this thesis is the Salt River Project (SRP) Digital Library (tDAR id:57630). According to its website, SRP is a not-for-profit public power and utilities company in central Arizona that serves over two million people. SRP has been involved in the development of the greater Phoenix area since it was founded in 1903. The company involves

CRM in their projects to mitigate the impact of their development of infrastructure on the archaeological record. This practice, legally mandated through Section 106 of the NHPA, has resulted in a vast collection of CRM gray literature reports. These reports and their associated data and images were originally held in analog form in the SRP office until it was relocated. The files were then digitized and uploaded to the SRP Digital Library collection in tDAR. This shift to a digital repository made it possible to better manage the large collection and ensure the preservation of the resources within it. All files in the SRP Digital Library were originally restricted to access solely by the administrator, an archaeologist working for SRP. After Dan Garcia assumed the sole archaeologist position in 2019, he sought to change the role of the SRP Digital Library from an in-house-only library to a resource for anyone to learn more about the archaeology of the Salt River Valley. CRM professionals, academic researchers, and the public were thus able to request access to certain files within the collection.

The original canal systems that the SRP expanded upon for modern day water and power supply were created by the Hohokam (Salt River Project n.d.). The Hohokam built and maintained over 500 miles of canals in the Salt River Valley that were integral to life and agriculture in the Hohokam community for over 1,000 years (Cordell and McBrinn 2012:38,164; Hill et al. 2015:610). It is estimated that just one major trunkline of the impressive canal systems took nearly one million person-days to construct, not including any secondary lines or subsequent repairs and maintenance (Fish and Fish 2008:5). It is difficult to pinpoint the exact reason for the Hohokam collapse, but investigations suggest a combination of social conflict and both human and naturally induced ecological problems (Hill et al. 2015:648). However, the Hohokam did not disappear. The O'odham living in the Salt River Valley today identify

themselves and their descendants as Huhugam. Their heritage, including their language, economy, ecology and other traditions, supports this connection (Hill 2019:2).

Considering the complex archaeological and cultural history of the region, access to the resources within the SRP Digital Library is regulated. Garcia singlehandedly manages all requests and ensures that users looking to access certain files have a legitimate reason to do so. This is important because the reports often contain specific locational information and images that stakeholders, including descendant communities, do not want publicized. As discussed in greater detail in Chapter 4, the SRP Digital Library includes public versions of some resources in which these sensitive data have been redacted.

Chapter 3: Methods

Methodology Overview

The aim of this thesis is to better understand how archaeologists can effectively and ethically manage access to archaeological data to allow for dynamic reuse. The methods for doing so include examining the structure, content, accessibility, and reuse of two databases from digital archaeological data repositories. These four characteristics were chosen to illustrate how different projects can have differing qualities yet still allow for productive interactions with their data. The order in which the four characteristics are addressed is intentional and important for the analysis as they build upon each other. Reuse is dependent on the accessibility of the data, accessibility is dependent on the sensitivity and type of content, and the content is dependent on the structure of the database.

As discussed in depth in Chapter 2, the case studies are the Mimbres Pottery Images Digital Database (MimPIDD) and the Salt River Project (SRP) Digital Library. MimPIDD and the SRP Digital Library were chosen in part because they have certain characteristics that differ from each other and that allow for meaningful conclusions to be drawn about how those characteristics impact the reuse/potential for reuse of the databases. Moreover, these two case studies were selected for this research because they are both well-established and commonly accessed resources on tDAR. Their popularity ensures that there will be ample view and download data to work with. It also means that there are likely to be more instances of reuse than there would be for newer, less popular databases.

Structure

The structure is the foundational aspect of digital data management upon which every other aspect is built. In this context, the structure of a database refers to the:

- Creator(s) and administrator(s) of the database,
- Reason for creating the database,
- Organizational framework of the database itself, and
- Organizational framework of the platform that houses the database.

The creators and administrators of the case studies can provide insight into how and why each database was created. In the discussion that follows I consider the institution(s) that played a role in the creation of each database as well as the creator(s) and administrator(s)'s role within the institution(s). I will also discuss when the databases were created and how the leadership structure has changed over time. These characteristics can be identified and described based on metadata, internal documentation, and discussions with the creators and administrators.

To analyze the structure of the two case studies I will be looking at the platforms used to house the databases as well as the motivations of the creator(s) and administrator(s) of the projects. This includes consideration of the type of resource each database has been designated as in tDAR (see Figure 3.1). Collections and projects are two of the many resource types in tDAR. Both are organizational resources that allow creators and administrators to customize groups and hierarchies of data resources within them. Collections have a further benefit of giving the administrators control over which users can view and download specific files (Digital Antiquity 2021). Understanding these structural aspects of each case study will help contextualize the reasoning behind their content, accessibility, and reuse.

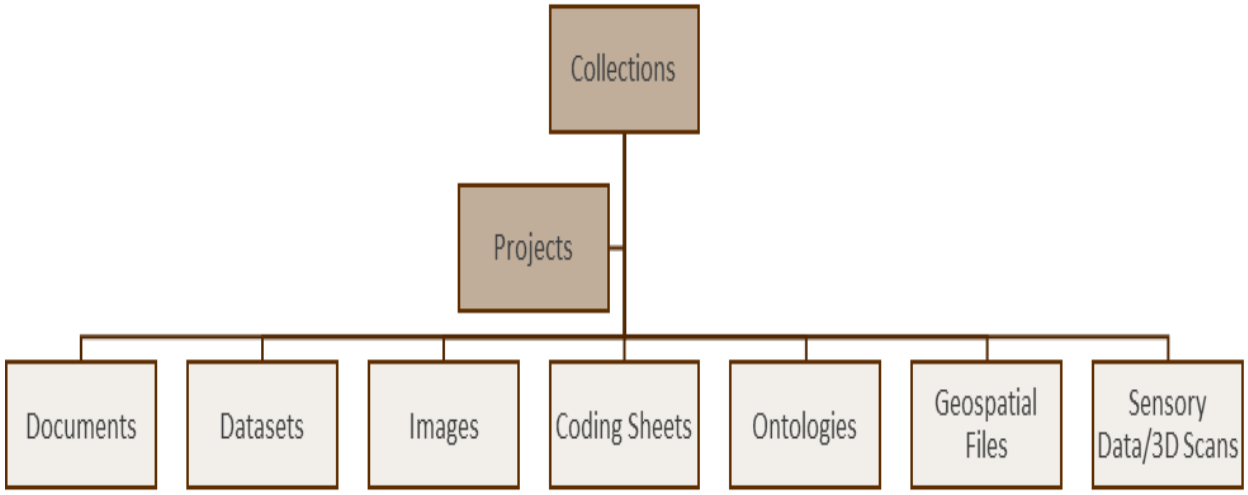

Figure 3.1. Resource types in tDAR.

Content

It is also important to investigate the content of the projects as this is where the capacity for reuse is most obvious. For the purposes of this study, the content of a database includes the:

- Types of resources (see Figure 3.1),
- Number of resources,
- Level of metadata associated with each resource and the project as a whole, and
- Cultural background of the materials.

All resource types have varying potential for reuse. For example, an image with associated attributes like the resources in MimPIDD can be reused differently than a gray literature CRM report like those in the SRP Digital Library. Analyzing the content also includes understanding the archaeological cultures involved and how they affect ethical considerations regarding what constitutes sensitive information. This is where both the FAIR and CARE Principles become key. In this section of the analysis, I will examine the qualities of the case

studies that fall under these theoretical principles and how they determine the appropriate level of accessibility for resources in the databases.

Accessibility

The accessibility section of the analysis focuses on interactions with the resources in each case study. For this aspect of the research, I will be looking at:

- How access to culturally sensitive data is managed by the administrators,
- The number of views of the database overall,
- The number of views for a sample of resources, and
- The number of downloads for a sample of resources.

This section is meant to illustrate how different levels of access to sensitive data are managed by the administrators and how this affects the quantifiable interactions with the resources, if at all. The quantifiable interactions are included in the analysis to better understand the accessibility, in terms of reach and influence, of each database. This will become even more pertinent when discussing instances of reuse.

The expected results for the analysis of the number of views and downloads per file type are as follows:

- Views and downloads for MimPIDD's access-restricted *research* dataset and the SRP Digital Library's *confidential* files could indicate how professional and academic audiences are engaging with the materials.
- Views and downloads for MimPIDD's *public* dataset and the SRP Digital Library's *redacted* files could suggest how public audience are engaging with the materials

and/or how professional and academic audiences who do not need detailed spatial or burial data are engaging with the materials.

Reuse

As discussed in previous chapters, it is difficult to measure the reuse of digital data. As discussed by Huggett (2018:95), searching for DOI's often only leads back to the original resources, not all instances of reuse are published, and those that are published are not always cited fully. We must continue to search for effective ways of identifying all instances of reuse. In the meantime, user surveys remain the best tool for acquiring such information.

The aim of this section is to better understand the potential for data reuse and the overall success of MimPIDD and the SRP Digital Library in facilitating the reuse of data. For this aspect of the analysis, I will look at:

- The number of times data from the case studies have been reused,
- The accessibility of the data reused (e.g., publicly accessible vs. requires permission),
- Who is reusing the data, and
- The context of reuse (e.g., academic, professional, educational).

A qualitative analysis of MimPIDD data reuse will also be performed. This includes analyzing three examples of reuse from three categories: reuse in professional publications and presentations by MimPIDD board members, educational uses organized by MimPIDD board members, and various uses by others given access to the full research version of the database. The examples were chosen from a list of reuses provided by the MimPIDD administrators. The

analysis will include a discussion on who reused the data, the context of reuse, how the data from MimPIDD was utilized, and how the structure, content, and accessibility of the database, as outlined in the previous sections, contributed to the quality of reuse.

The expected results for the reuse analysis are less straightforward due to the experimental methodology. Some basic assumptions regarding the type of data reuse are as follows:

- Academic reuse could suggest that collection, management, and metadata recordation are accurate and meet required standards of use in accordance with the FAIR and CARE Principles.
- Public reuse could suggest that data are accessible both physically and in terms of the language used to describe the data.
- Little-to-no examples of reuse could suggest that the data do not meet the technical standards necessary for reuse and/or that the project is too new to have seen instances of reuse and/or that prospective users are unaware of the potential of reuse.

Data Collection

Much of the data used in this thesis were accessible via a basic tDAR login. tDAR's data accessibility framework (see Figure 2.3) ensures that up-to-date metadata on view and download metrics are available to all users logged into tDAR. As shown in Figure 3.2, these data appear at the bottom of every resource page. Because these metrics only show the statistics at the time of inspection, I had to gather long-term data from other sources. tDAR's website also contains a blog wherein updates on the platform and other news items are uploaded. Yearly summaries on

the overall use of tDAR were uploaded to this blog between 2014 and 2017 by Adam Brin (2015, 2016, 2017, 2018). These summaries include statistics on the most viewed and downloaded resources and collections in tDAR as well as the total number of resources in the repository.

Resource-specific year-by-year view and download data were also graciously provided by Rachel Fernandez, Digital Preservation Program Manager at Digital Antiquity. Fernandez, as well as Christopher Nicholson, Director of Digital Antiquity, met with me over zoom and provided additional anecdotal information on tDAR and the two case studies. As noted previously, data on reuse were hard to obtain due to tracking difficulties. Will Russell, the current administrator of MimPIDD, shared a running list of known instances of reuse compiled by the MimPIDD board. Dan Garcia, Senior Cultural Resource Management Specialist at the Salt River Project, shared anecdotal instances of reuse of data from the SRP Digital Library as well as contact information of potential reusers. This research would not have been possible without the generous help of these individuals.

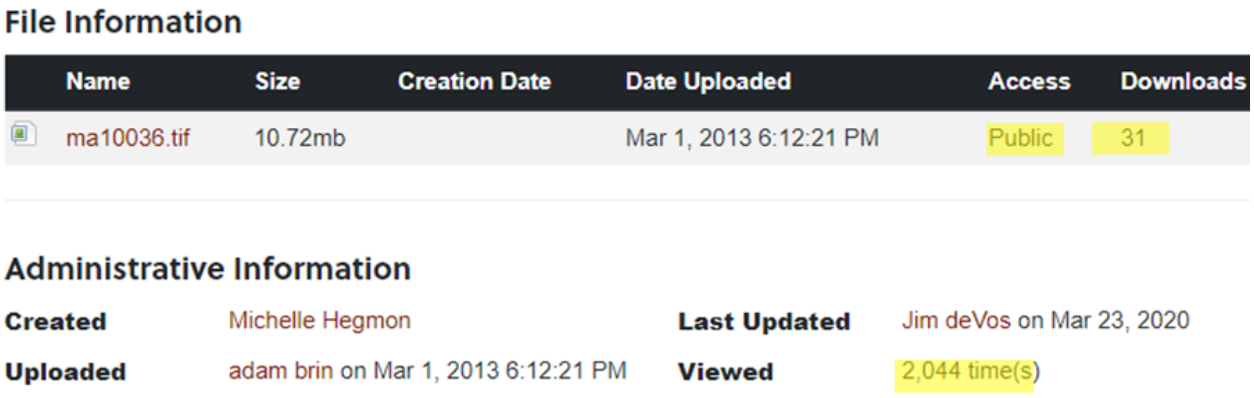

Figure 3.2. Example of view and download metadata for a MimPIDD file in tDAR.

Chapter 4: Analysis and Results

Mimbres Pottery Images Digital Database (MimPIDD)

Structure

MimPIDD started to take form in 2003 under the direction of Michelle Hegmon and Steve LeBlanc, but the concept of a cumulative database of Mimbres pottery had been in the works long before then (Hegmon et al. 2017:2). One of the main goals of the database is to make large-scale comparative analyses of otherwise dispersed data on Mimbres pottery possible. Hegmon has been affiliated with Arizona State University, the institution that created and manages tDAR, since before tDAR's inception. This close connection could explain the success of the project. Until the summer of 2022, Hegmon was the chair of the MimPIDD board. The position has since been assumed by Will Russell, Historic Preservation Specialist with the Arizona Department of Transportation.

MimPIDD is a collection of data housed in tDAR. As discussed in Chapter 3, collections are a type of organizational resource in tDAR that are used to group resources together. In the case of MimPIDD, being set up as a collection not only allows for structured organization of multiple resource types within the database, but it serves an even more important purpose of allowing the administrators to differentiate access to resources. MimPIDD is made up of two databases, one that is accessible to all members of the public, and a second version that has restricted access. This differentiated access is discussed further in the *Accessibility* subsection below.

Content

MimPIDD contains datasets, documents, and images pertaining to over 9,000 Mimbres pottery vessels that are scattered worldwide due to looting and artifact trafficking. These 9,000+ vessels represent the majority of known Mimbres pottery (Hegmon et al. 2017:2). The two versions of the database are different in three important ways: their content, their sources, and their accessibility, the last of which will be discussed in detail in the *Accessibility* subsection below. As of August 10, 2022, the public version of MimPIDD contained 2,188 resources including two datasets, 15 documents, 2,168 images, and three projects. The research version contains the same two datasets, three documents (two of which are duplicates from the public version), and an additional 7,395 images (Table 4.1).

The two datasets, the *Mimbres Pottery Database (Public)* (tDAR id:22070) and the *MimPIDD Research Database for Download* (tDAR id:381505), contain the same images and attribute data available on the MimPIDD tDAR page, but they are tabulated in Microsoft Excel files and the research dataset has additional rows for the images that are in the research version of the online database but not the public version. While the public and research versions of MimPIDD have the same two datasets listed as resources in the collections, the *Research Database for Download* is only listed (i.e., it is not accessible) in the public version; it can only be downloaded upon request and if the administrator approves access to the research version of MimPIDD. Since the research version of MimPIDD contains thousands of additional images, there are thousands of additional rows of attribute data for those vessels in the research dataset.

Considering MimPIDD's primary goal of bringing together as much data as possible on Mimbres pottery, the resources in the database come from various sources. The public version of MimPIDD contains data sourced from universities, museums, and other public-facing

organizations that have granted permission for MimPIDD to make their data available to the public. In contrast, the research version includes data from private collections and other institutions that did not choose to grant full, unrestricted access to the material as well as artifacts that have questionable provenance histories and connections to looted sites and burials. This is why there are fewer images in the public database. In addition to the sources of data, the attributes of the two versions of MimPIDD differ. Both datasets contain the same base-level attributes such as archaeological sites, vessel diameters, and designs present on vessels. The research version, however, has ten additional attributes that relate to specific locational and burial information. Considering that many Mimbres vessels came from burial contexts, it is important to protect associated data that are sensitive (Hegmon et al. 2017:2).

Table 4.1. Downloadable resources per version of MimPIDD.

Public Version	Research Version
Public Dataset	Public Dataset
Ethics and Permission to Access MimPIDD	Ethics and Permission to Access MimPIDD
MimPIDD Users' Guide	MimPIDD Users' Guide
Education Resources	Education Resources
2,168 Images	9,583 Images
	Hegmon and Russell 2013
	Research Database

Thirteen of the 15 documents in the public version of MimPIDD are related to three projects: *College/University Education Resources* (tDAR id:455444), *Grades 4-12 Education Resources* (tDAR id:455448), and *K-3 Education Resources* (tDAR id:455439). These projects and the documents within them include grade-appropriate lessons that teach students about data coding and artifact identification. The other two documents—which are included in both the public and research versions of MimPIDD—are the *Ethics and Permission to Access MimPIDD* (tDAR id:381421), which contains information on how to get access to the research version of the database and the ethical implications of it, and its *MimPIDD Users' Guide* (tDAR id:455490), which contains a history of MimPIDD, the same ethics and permission to access MimPIDD information, and coding sheets for the attributes in the datasets. The research version also contains an additional document (tDAR id:377853), a report written by Hegmon and Russell (2013) on identifying unique Mimbres artists using the data from MimPIDD.

Images are undeniably MimPIDD's largest component. Mimbres pottery is renowned for its artistic quality and the images display a wide variety of designs. The searchable database component of the collection allows users to search by keywords including general or specific figurative or geometric designs. Each image has a download link and is accompanied by the same attribute data found in the datasets.

Accessibility

MimPIDD uses what I refer to as a double-database approach to manage access to resources within the collection. The public version of MimPIDD can be downloaded by anyone with a free and easy to set up tDAR login. This same public dataset is duplicated in MimPIDD's online searchable database. The research version of MimPIDD, on the other hand, requires

approval to access. By having two databases with differentiated access, the administrators can ensure that public accessibility goals are being met while also protecting more sensitive data by only making them available to vetted users.

As mentioned in the *Content* section above, there are specific directions that tDAR users must follow if they wish to gain access to the research version of MimPIDD. There is a resource called *Ethics and Permission to Access MimPIDD* within the MimPIDD collection. Detailed instructions for how to access the full research database are listed alongside a discussion of ethical concerns surrounding the materials in MimPIDD. To gain access to the research database, users must send a signed request to the MimPIDD administrator. In the request the user must:

- describe their research,
- agree that they will not share the materials with others,
- make clear that they understand the problems with non-authentic designs and the need to be cautious about analyses that include non-provenienced vessels,
- indicate that they understand that they cannot publish the MimPIDD photographs without permission from the owners of the vessel, and
- acknowledge the ethical problems with publishing material in private collections.

Views per Dataset

Figure 4.1 displays the total number of views for the research and public versions of the MimPIDD datasets as of October 1, 2022. The views represent the total number of times that tDAR users visited the specific webpage of the file. These webpages display the associated metadata and a download link for the file rather than the data themselves. Thus, the number of

views per dataset does not necessarily indicate much about data access and reuse because anyone can go online and view any listed page without seeing any of the actual data.

These data show that tDAR users have viewed MimPIDD's research dataset 3,168 times and viewed the public dataset 3,954 times. This difference of approximately 800 views suggests that the public dataset has a greater reach and/or has generated more interest than the research version. The previously discussed nature of these data makes it difficult to draw more meaningful conclusions from this comparison.

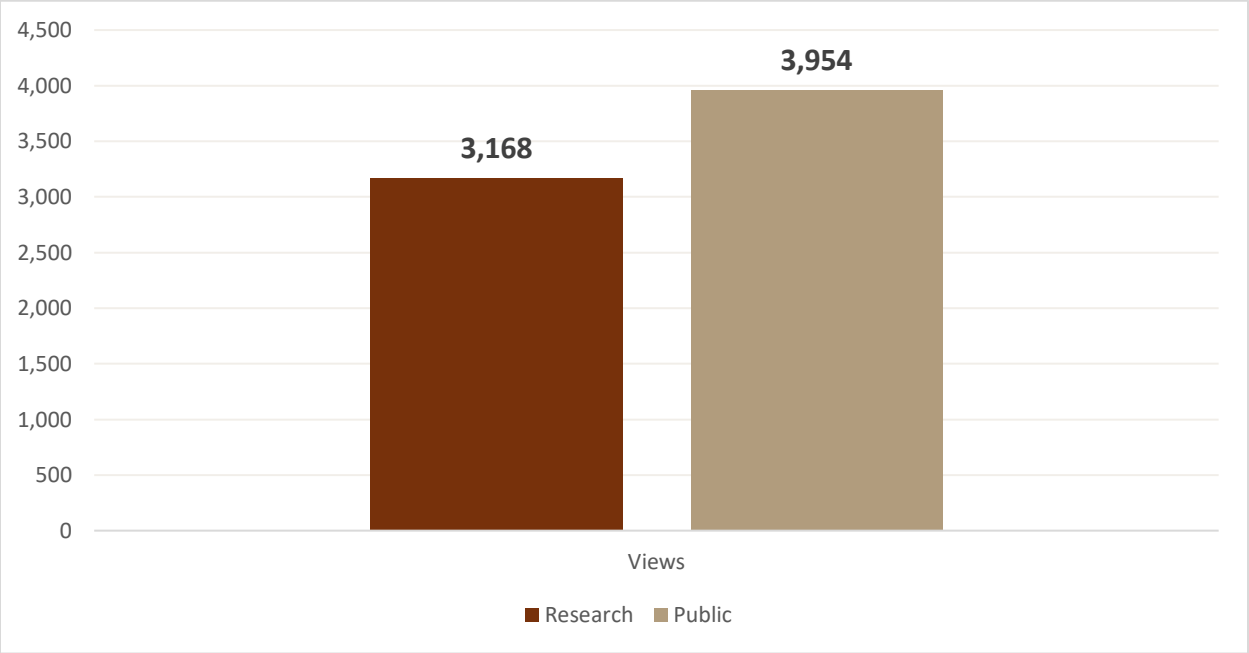

Figure 4.1. Total views of MimPIDD's research and public datasets as of Oct. 1, 2022.

Downloads per Dataset

Figure 4.2 shows the total number of downloads for the public and research versions of the MimPIDD datasets as of October 1, 2022. Considering that anyone with a tDAR login can

download the public dataset file, these downloads could represent members of the public interested in the tabular data and/or archaeologists who simply do not need the burial or locational information included in the research version of the dataset for their analyses.

According to the file metadata, tDAR users downloaded the public dataset 161 times since it was first uploaded to tDAR in 2013. Even without knowing who downloaded the dataset, the high number of downloads is quite impressive given that users can easily search through the data in the online database. Although the number of downloads of the research dataset is only approximately one-third that of the public version, the 53 downloads are still substantial. They suggest that researchers are finding and using this database and thus that tDAR and MimPIDD's goals are being realized.

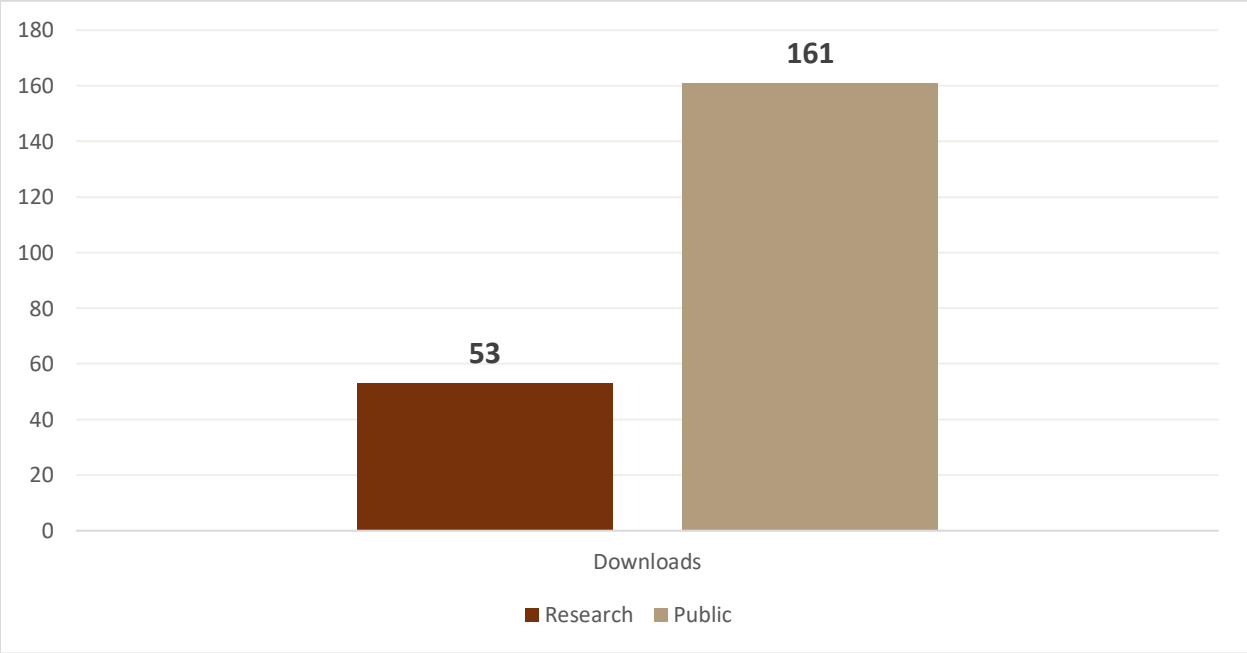

Figure 4.2. Total downloads of MimPIDD's research and public datasets as of Oct. 1, 2022.

Overall Trends

tDAR published annual reports of their usage statistics on their blog, which is accessible on their website. The posts included data on the most viewed and downloaded resources as well as the overall trends in their growth as a platform. Unfortunately, tDAR has not published summative blog posts since January 2018. Nevertheless, the available data—which span the period from 2014 to 2017—can shed light on not only the development of the entire repository but the trends of the projects within it, including MimPIDD. Given that MimPIDD was first added to tDAR in 2013, these data show how the interactions with the database have grown since not long after its inception. MimPIDD was the second and fourth most viewed resource in tDAR in 2014 and 2015 respectively (Brin 2015, 2016), and the most viewed resource in both 2016 and 2017.

The public dataset was similarly successful in the broader tDAR context. Throughout the four-year period, the dataset was one of the top thirteen most downloaded resources in tDAR (Brin 2015, 2016, 2017, 2018). To put these statistics into perspective, tDAR was home to over 400,000 resources by the end of 2017 (Brin 2018). As shown in Figure 4.3, additional data provided by Rachel Fernandez (personal communication 2022) suggest that the number of downloads per year has been declining for both the public and research datasets. There are, however, occasional peaks (e.g., the research dataset in 2018) and each dataset has been downloaded at least once per year. This is still impressive as it points to consistent traction and engagement with MimPIDD over the last decade. Together, these data suggest that the goal of increasing public access to data via MimPIDD has been successful, and the consistency of downloads through time suggest that the data continue to be relevant.

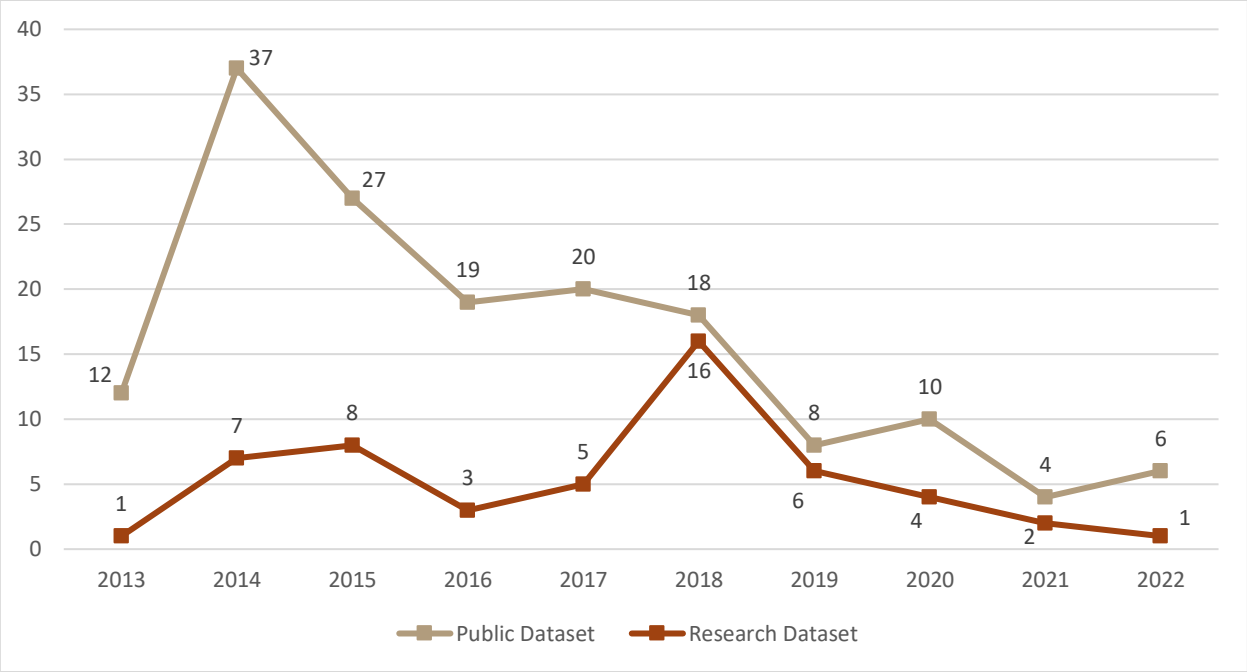

Figure 4.3. MimPIDD public and research dataset downloads per year from 2013 to October 1, 2022.

Reuse

As discussed in Chapter 2, there is little methodology in place to gather data on and interpret the reuse of data from digital repositories. Michelle Hegmon and Will Russell, the current chair of MimPIDD, have been curating a list of how the data in MimPIDD has been reused. The list was first compiled in March of 2022, and it is continually added to as more uses become known. Russell graciously shared the list with me to use for this section of the thesis.

The data were initially collected by MimPIDD board members and were later supplemented by responses to an email sent to individuals who were granted access to the research version of the database. With this in mind, it is important to note that there are undoubtedly more instances of reuse that remain unknown, especially by users of the public version of the dataset who were not included in the survey. While ambitious, a more widespread

survey asking all users who downloaded any resource in the MimPIDD collection how the resources were used would provide a more thorough view of the collection's reuse. The following analysis focuses on the known instances of reuse, including this thesis, as of August 2022. There are surely more instances of reuse that have not been taken into account because they have not been communicated to the administrators. Additionally, there are at least three other known instances of reuse that are still in the publication process, and thus were not included in the quantitative analysis.

There are a total of 54 recorded instances of reuse of data from MimPIDD. These may be split into three major categories. As shown in Table 4.2, there are 29 instances of reuse in the form of professional presentations and publications by MimPIDD board members, three examples of educational uses organized by MimPIDD board members, and 22 instances of reuse (mainly professional presentations and publications) by other individuals who were given access to the full research version.

Table 4.2. Instances of MimPIDD data reuse by category.

Category	Instances of Reuse
Presentations and publications by MimPIDD board members	29
Educational uses by MimPIDD board members	3
Various uses by others given full access	22
Total	**54**

The most noticeable characteristic of the MimPIDD reuse data is that the majority of known instances of reuse are by MimPIDD board members. Fifty-four percent of all recorded instances of reuse are professional publications and presentations that least one of the six MimPIDD board members was involved in; an additional 5 percent comprises educational uses organized by the board members (Figure 4.4). The remaining 41 percent of cases involve data reuse by other individuals given access to the full research version of MimPIDD. However, these data could be biased considering that motivation to identify uses is likely higher among researchers involved with the project, and the email survey was likely not completed by all users to whom it was sent.

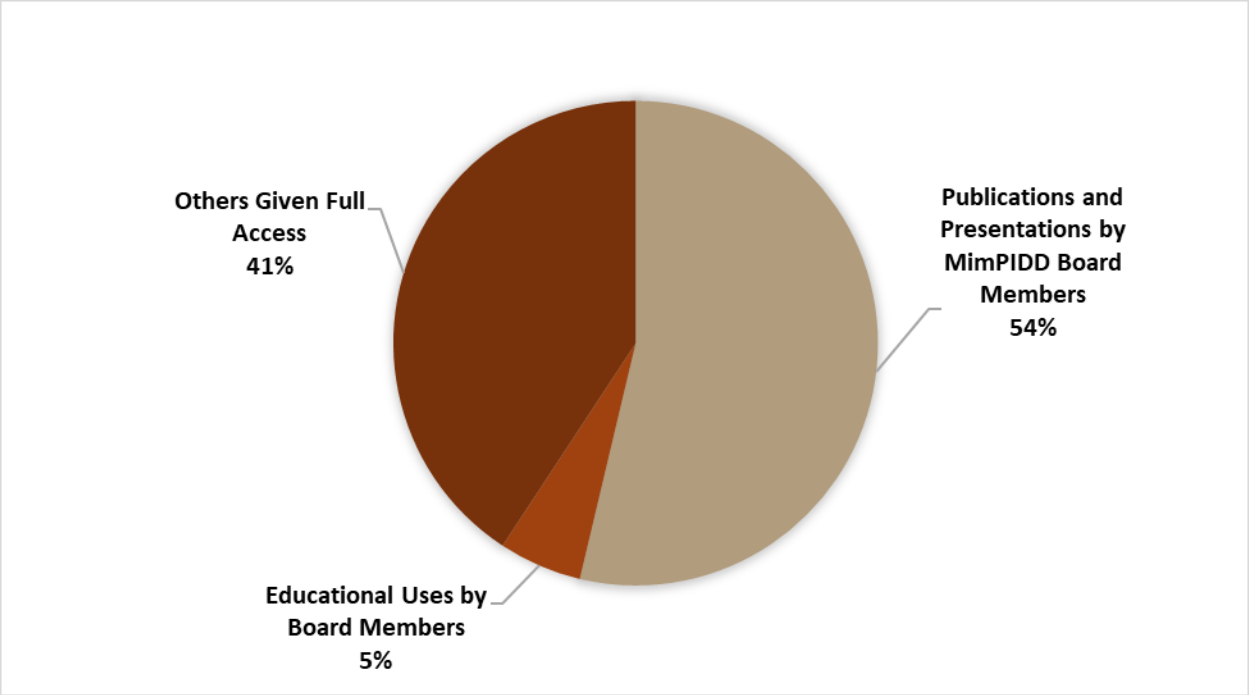

Figure 4.4. Frequencies of different categories of MimPIDD data reuse.

Of the 54 total instances of reuse, there are 26 unique first authors and 46 unique authors overall. When broken into the same categories discussed above, there are nine unique first

authors from the MimPIDD board, three unique first authors who used the data for educational purposes, and 13 unique first authors among the other individuals granted access to the research version. Moreover, two of the nine first authors in the first category are currently on the MimPIDD board, while the other seven instances of reuse in that category are either by previous board members or were co-authored by board members.

Considering that the document used to record the instances of reuse by the MimPIDD administrators contains citations, the frequency of reuse can be easily tracked year by year. As shown in Figures 4.5 and 4.6, there has been a fair amount of variation in the frequency of reuse per year. One might assume that the instances of reuse per year would generally follow the trends evident in the views and downloads per year. However, as Huggett (2018:94) mentions, views only suggest that the materials are being accessed while downloads simply point to the potential for reuse since individuals downloading the data may not use them, and if they do, there can be quite a large gap in time from the inception of a research project to its eventual publication.

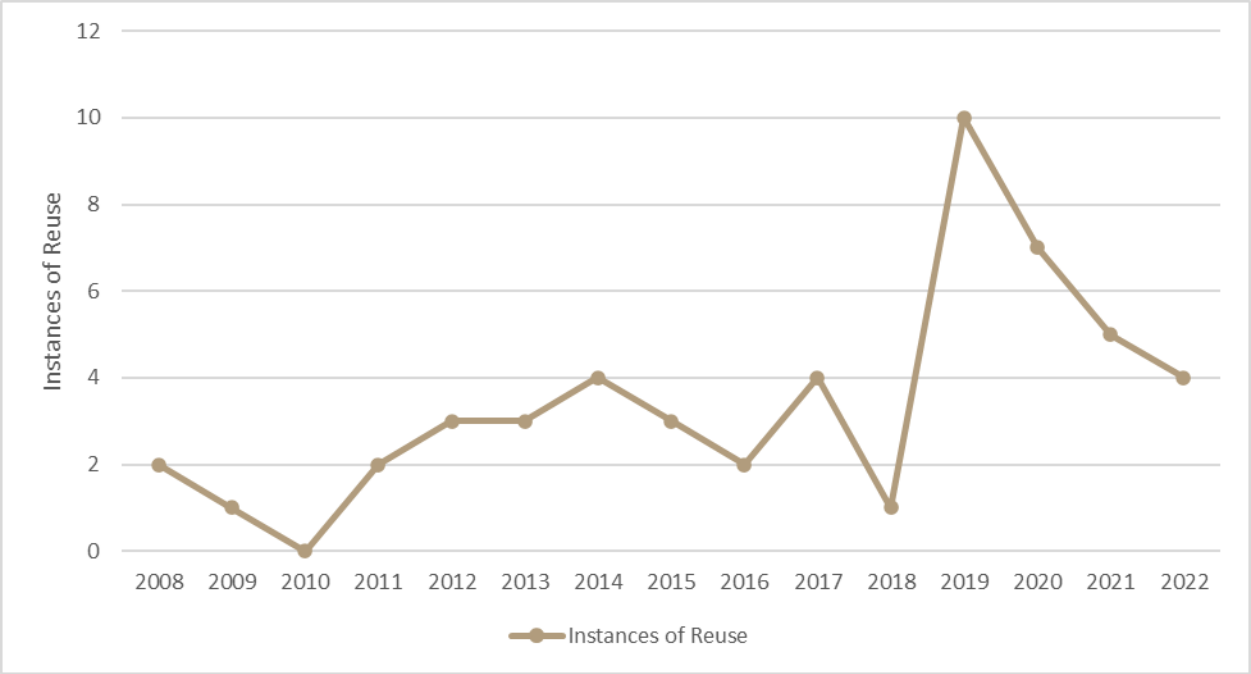

Figure 4.5. Overall instances of MimPIDD data reuse over time.

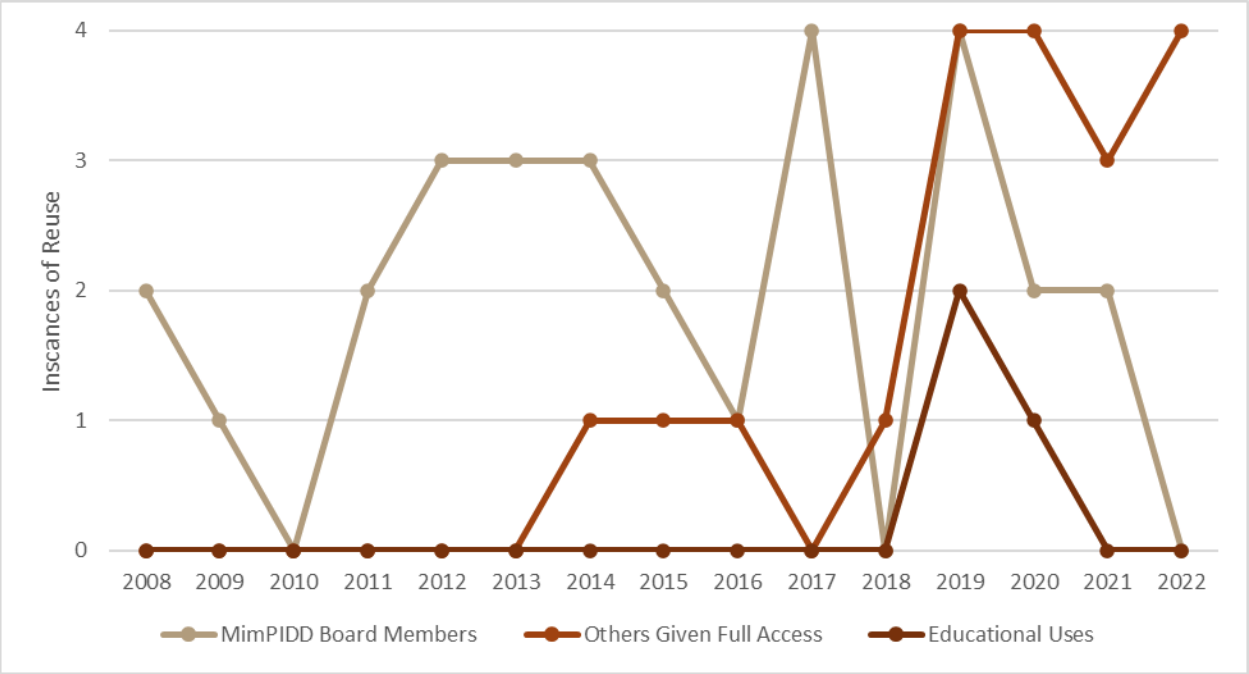

Figure 4.6. Instances of MimPIDD data reuse by category.

Qualitative Analysis of Reuse

To better understand the nature of MimPIDD data reuse, this section will explore an example of reuse from each of the three categories discussed above (professional publications and presentations by MimPIDD board members, educational uses organized by MimPIDD board members, and various uses by others given access to the full research version). The specific examples were chosen to show the variety of ways in which the data in MimPIDD have been reused in each of these three contexts.

Professional Publication by MimPIDD Board Members

One example of reuse is an article entitled "The Social Significance of Mimbres Painted Pottery in the U.S. Southwest" (Hegmon et al. 2021) that was published in *American Antiquity* and can be accessed online through Cambridge Core. Two of the authors of this article, Michelle Hegmon and Will G. Russell, are current members of the MimPIDD board; Russell took over the role of chair from Hegmon in the summer of 2022. Combined, Hegmon and Russell have been involved in 31 of the 54 total recorded instances of reuse, including every one of the 29 instances of reuse by MimPIDD board members and two of the three educational uses.

The article was electronically published by Cambridge University Press in August 2020. Cambridge Core tracks the access metrics for its resources. As of September 27, 2022, the abstract had been viewed 1,893 times, the full text had been downloaded 844 times, and the article had been cited twice (Cambridge Core 2020). It is important to note that this article is not open access. The full text can be downloaded for free with an institutional affiliation login that grants access to the resource or for a $25 fee.

In this article, Hegmon et al. use the attribute data and photographs available in the full research dataset in MimPIDD to analyze how variations in the designs on Mimbres Black-on-White Style III bowls were used to signal social acceptance and identity (Hegmon et al. 2021:23,39). Using data on 1,056 bowls in MimPIDD, they compare the proportions of representational and geometric designs along with the presence or absence of kill holes and the frequencies of bowls from burial contexts. Their analysis of this representative sample from three well-excavated sites indicates that the bowls were used in similar ways, yet geometric designs were nearly twice as common as representational designs (Hegmon et al. 2021:27).

The researchers then build upon that idea to argue that Mimbres pottery is bounded in three ways: by the use of distinctive designs, based on the region in which the pottery is commonly found, and based on the period when there was an abrupt decline in population at Classic Mimbres sites and Mimbres pottery production halted (Hegmon et al. 2021:29). They identify three dimensions of variation seen in Mimbres pottery designs using data from MimPIDD. These include identification of specific individuals that used unique design styles, minimal intraregional variation, and the social significance of the spatially bounded Mimbres pottery (Hegmon et al. 2021:29-34). The analysis that led to the identification of minimal interregional variation centered on the distribution of geometric designs of bowls from four sites: Cameron Creek, Mattocks, Pruitt, and Swarts. The MimPIDD-sourced data revealed no major differences in the distribution of certain designs, suggesting that the same general suite of designs was present at every site (Hegmon et al. 2021:33). Further, the spatial boundedness of Mimbres pottery was supported by a representational sample of 361 bowls from Swarts, Mattocks, and NAN Ranch that "confirms the existence of consistently used rules of design" (Hegmon et al. 2021:34).

The article has an associated project in tDAR (tDAR id:455455) that contains coding sheets and datasets with the data extracted from MimPIDD that were used for the analysis. This project and the resources within it exemplify the cyclical nature of data publication and reuse as these data tables, while not containing *new* data per se, can themselves be reused. Considering that they are housed in tDAR, their access metrics, such as views and downloads, can also be investigated, as demonstrated by this thesis.

Hegmon et al.'s study was enabled by MimPIDD's structure, content, and accessibility. The database was designed specifically for use in large-scale analyses of this kind and the detailed attribute data and multitude of images in the database allowed the authors to perform comprehensive analyses on designs and their distribution. The authors already had access to the full research dataset as members of the MimPIDD board. This article exemplifies how MimPIDD data can be (re)used in academic research.

Educational Use Organized by MimPIDD Board Members

As discussed above, MimPIDD also contains a set of K-through-12 educational materials (tDAR id:69827). These resources were created by April Kamp-Whittaker and Michelle Hegmon in 2020 as part of a National Endowment for the Humanities (NEH) grant project called "From Library to Laboratory: Developing Tools to Enhance the Use of Digital Archaeological and Other Humanities Collections." This educational program contains resources relating to data coding, artifact identification, and artistic design.

The educational resources are organized into three projects within MimPIDD: K-3 Education Resources, Grades 4-12 Educational Resources, and College/University Education Resources. The K-3 project contains four resources, the 4-12 project has five, and the

College/University project has three. There is an additional lone-standing resource that includes instructions on how to search tDAR. The resources in each project include both educational activities as well as instructions for teachers. All materials in the education resource projects are publicly available for download through MimPIDD. Table 4.3 shows the number of resources, as well as the total numbers of resource views and downloads, for each project. Notably, the 4-12 and College/University educational materials were each downloaded approximately twice as often as the K-3 educational materials.

Table 4.3. Views and downloads of MimPIDD's educational projects as of Oct. 1, 2022.

Project	**Number of Resources**	**Total number of resource downloads**	**Total number of resource views**
K-3 Education Resources	4	20	1,519
4-12 Education Resources	5	41	2,004
College/University Education Resources	3	43	1,332
Totals	12	104	4,855

The lessons reference materials within the public version of the MimPIDD database. The K-3 project contains simple coding activities that introduce young students to Mimbres archaeology and how to identify general figurative characteristics on Mimbres bowls. Both the 4-12 and the College/University activities require students to search MimPIDD and download specific images. This could account for many of the views and downloads of specific resources within MimPIDD.

The 4-12 project includes a reading covering the history of Mimbres archaeology and activities that teach students how to search through MimPIDD for images of specific classes of animals and find all images of a particular animal from a specific site. The College/University activities introduce students to data coding and how to search for resources in tDAR. The lessons instruct students to search for specific Mimbres bowls in MimPIDD and record descriptions of their attributes and discuss their methods for searching for, identifying, and analyzing pottery designs.

The search feature in MimPIDD plays a major role in making these educational resources as impactful as they are. Students only need basic attribute data to complete the activities created by Kamp-Whittaker and Hegmon. Thus, the fact that attribute data and images of many vessels are publicly available in MimPIDD allows for public-facing data access and reuse like these education resources to reach classrooms worldwide.

Use by Others Given Full Access to MimPIDD

The final example of reuse is a blogpost written by then-undergraduate student Lindsay Shepard entitled "Using a Multidisciplinary Approach to Interpret Artifacts." This blogpost was published in 2015 in Archaeology Southwest's publicly accessible *Preservation Archaeology Blog*. Archaeology Southwest is a non-profit organization that practices preservation archaeology, "a holistic and conservation-based approach to exploring and protecting heritage places while honoring their diverse values" (Archaeology Southwest 2022).

Shepard's blogpost discusses how incorporating a multidisciplinary approach to artifact interpretation can lead to new, helpful perspectives on what artifacts represented to the people who used them. The blogpost focuses on figurative designs on Mimbres pottery vessels,

specifically a Mimbres Late Style II to Middle Style III bowl from MimPIDD with a mobile-like figurative design. Shepard (2015) suggests that the mobile-like apparatus depicted on the Mimbres bowl could represent the same type of object as composite wooden artifacts found in the region. Shepard also discusses the importance of ethnographic analogy in decoding the meaning of objects represented on Mimbres pottery. She argues that archaeologists can better understand the function of items depicted in Mimbres pottery by engaging in conversations with Indigenous groups about their current cultural practices. The approach that she advocates in this blogpost is very similar to that posed by the MimPIDD board: a multi-sourced, collaborative, and ethically conscious archaeology.

This blogpost is written for a public audience, thus achieving MimPIDD's goal of making Mimbres archaeology better understood and more easily accessed by the public. A photograph of the vessel Shepard analyzed is published in the blog, with permission from the institution that owns it, following MimPIDD's guidelines. While this instance of reuse does not use tabulated data like the first example discussed above, it shows that the content in MimPIDD can be used in a wide range of contexts and can serve a range of purposes and audiences.

Salt River Project Digital Library

Structure

The SRP Digital Library (tDAR id:57630) was first created in tDAR in 2017 when the Salt River Project was moving offices and subsequently had to downsize their physical storage capacity. SRP's archaeologist at the time scanned and uploaded the documents to tDAR to resolve the storage predicament and ensure the longevity of their large collection of gray literature and CRM reports pertaining to their service region in the Phoenix metro area.

At the time of its creation, the SRP Digital Library was intended to be used solely by the SRP archaeologist and there were no plans to expand access to people outside the company. The use of tDAR for private data management is not uncommon as the repository ensures data security and longevity (Digital Antiquity 2018a; McManamon et al. 2021:243). When Dan Garcia stepped into the SRP archaeologist role in 2019, he recognized the potential benefits of making the large library accessible to other CRM professionals, academic researchers, students, and the public (Dan Garcia, personal communication 2022). Garcia has since worked to improve the sharing of documents in the SRP Digital Library with individuals outside of SRP.

Content

The SRP Digital Library is a collection housed in tDAR. As of September 15, 2022, the SRP Digital Library contained a total of 776 resources, including ten datasets, 761 documents, three images, and two projects. Nearly all the documents in the SRP Digital Library are gray literature and technical reports from CRM investigations. The documents are OCR (optical character recognition) readable, meaning that the text within the files is machine readable, improving their findability. The three image resources contain a total of 123 images, all from the same survey project. These photographs and an associated report make up one of the two projects in the SRP Digital Library. The other project contains a document, the final report of the archaeological investigation, and all 10 datasets (appendices from the report) from a survey for the Arizona Department of Transportation.

The data in this collection pertain to archaeological investigations in the Salt River Valley. As discussed in Chapter 2, this region was occupied by the Hohokam from 1500 B.C. to A.D. 1450. “Hohokam” is an archaeological term referring to this specific period while the term

"Huhugam" is more inclusive and includes the O'odham ancestors living in the region after 1450 (Archaeology Southwest 2017). The original SRP water network is based on the canal systems built by Hohokam groups to irrigate their farms (Archaeology Southwest 2017; The Salt River Project n.d.).

Nearly all files in the collection are marked as "Confidential." When uploading files to tDAR, users have the option to mark a file as "Confidential," making only the file metadata visible to the public while access to the document remains restricted to administrator-delegated individuals (Digital Antiquity 2013). Along with the confidential file, many resources include publicly accessible duplicate versions that have had all sensitive images and locational information redacted. The redacted versions of the resources within the SRP Digital Library were created as part of the Digital Archive of Huhugam Archaeology (DAHA) (tDAR id:30428). DAHA is a NEH-funded project undertaken by the Amerind Museum and Digital Antiquity team starting in 2017 with the goal of creating a comprehensive archive of materials relating to Huhugam archaeology. Many of the over 2,000 resources in DAHA are derived from large collections of technical reports like those that make up the SRP Digital Library.

The SRP Digital Library is not a static database, but rather one that Garcia wishes to keep adding resources to (Dan Garcia, personal communication 2022). Recording metadata and redacting sensitive data are time-consuming so Garcia, as the sole archaeologist at SRP, has not been able to add more files to the collection recently. Nevertheless, the SRP Digital Library will continue to expand as funding and assistance become available.

Accessibility

The redacted resources in the SRP Digital Library are downloadable for anyone with a tDAR login. The redaction process allows cultural information in the SRP Digital Library and DAHA to be shared with the public without the risk of sensitive data and images being used to harm archaeological sites or ancestral communities. The confidential files within the collection, on the other hand, can be accessed through a resource-by-resource request-to-access system. All requests to access the confidential files are funneled through Garcia.

As discussed in Chapter 2, confidentiality is important to many archaeological datasets, especially those that contain sensitive locational and burial information. Access to confidential files in the SRP Digital Library can be requested directly through the tDAR page of the specific resources. As shown in Figure 4.7, users can send messages directly to the administrator of the resources explaining their research goals and why they would like access to the specific files. If Garcia considers the research angle to be appropriate and the case for access to the confidential resources is legitimate, he will grant access. Importantly, Garcia (personal communication 2022) is especially careful in granting access to resources with information on Native American tribal lands.

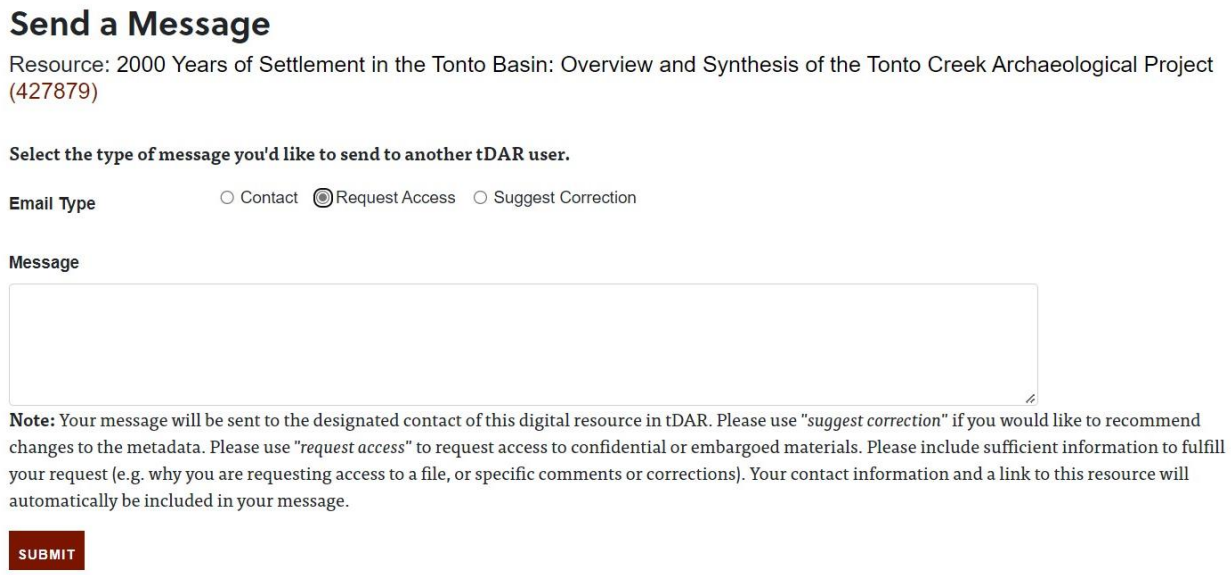

Figure 4.7. Request Access form on tDAR.

I have been told anecdotally by members of the tDAR team that the files in the SRP Digital Library are among the most requested resources in all of tDAR, with requests to access files being sent to Garcia almost every week (Christopher Nicholson and Rachel Fernandez, personal communication 2022). Approximately four months after talking to Nicholson and Fernandez, I met with Dan Garcia to discuss access and reuse of the SRP Digital Library. Garcia (personal communication 2022) added that while there has been a dramatic drop in the number of requests to access the SRP Digital Library resources since DAHA became available in 2021 and the redacted files have been published, but that he still receives frequent requests to access the restricted files. This is likely due to the easily accessible redacted resources including sufficient information for prospective researchers.

The four resources discussed below were chosen for this analysis because they contain both public/redacted and original confidential versions. Each resource was chosen by randomly

searching through the documents in the SRP Digital Library resource list and selecting resources with higher interaction (more views and downloads) and that also represent a variety of years, publishers, and CRM firms. It is important to emphasize that the redacted files are newer additions to the SRP Digital Library that were uploaded in the fall of 2021, while the original confidential files were added to tDAR in 2017. This needs to be considered when interpreting the statistical differentiation between the two resource types.

The first resource, "2000 Years of Settlement in the Tonto Basin" (tDAR id:427879), was published in 2004 by the Center of Desert Archaeology (renamed Archaeology Southwest in 2012). Arizona Department of Transportation (ADOT)-sponsored investigations, referred to as the Tonto Creek Archaeological Project (TCAP), resulted in this 366-page resource. The published report includes a summary of the large-scale archaeological investigations undertaken between 1992 and 1996 in preparation for an ADOT realignment. The project included the excavation of more than 100 structures and 300 burials at sites dating to the Middle Archaic, Classic, and Colonial periods (Clark and Vint 2004:1). The report synthesizes the results of the TCAP alongside other large-scale CRM investigations in the region.

The second resource, "Cultural Resources Survey for the SRP Palo Verde to Pinal West 500 kV Transmission Lines" (tDAR id:408577), was published in 2007 by Desert Archaeology Inc. This SRP-sponsored project was conducted in 2006 in preparation for the construction of two SRP transmission lines. The 109-page resource includes data from previous surveys in the area as well as the 2006 survey conducted by Desert Archaeology Inc. Eighteen total sites were identified along the corridor, with no sites or otherwise culturally significant properties being identified at the proposed substation site (Clark and Henderson 2007:79).

The third resource, "One Hundred Years of Archaeology at La Ciudad de Los Hornos" (tDAR id:427266), was published by Soil Systems Inc. in 1990. This project was funded by SRP in preparation for the construction of a new substation. Excavations of a Hohokam site, La Ciudad de los Hornos, were conducted in 1988 by Archaeological Consulting Services (ACS). The 139-page resource contains a discussion of the history of archaeological investigations in the region as well as the results of the 1988 survey that yielded 195 cultural features (Wilcox et al. 1990:iii).

The final resource, "Archaeological Investigations at Los Guanacos" (tDAR id:427775), was published by Northland Research Inc. in 1993. The investigations were sponsored by SRP in 1992 and resulted in the identification of 72 cultural features, including adobe-walled rooms, inhumations, and pit features (Howell 1993:3). The conclusions presented in the report shed light on how a failing canal system might have been one of the reasons the Hohokam inhabitants relocated (Howell 1993:180).

Views per File Type

It was expected that the access metrics for the SRP Digital Library would be different from those for MimPIDD because the materials consist mostly of gray literature. Moreover, unlike MimPIDD, the confidential and public versions of the SRP Digital Library are part of the same resource. This means that the views represent how *both* versions are accessed. Figure 4.8 shows the number of views for each of the sample resources, which are referred to by their tDAR ids. Overall, the files were viewed a comparable number of times, with the exception of "2000 Years of Settlement in the Tonto Basin" (tDAR id:427879). It is difficult to say for certain why this resource has been viewed more times than the others. It is possible that the resource could be

of interest to a wider range of users because of the large scope of the archaeological investigations (as discussed above) compared to the other files. Moreover, the higher number of views could also be due, at least in part, to the fact that resources in the SRP Digital Library are listed in alphabetical order, with numeric titles preceding the alphabetical titles. Because of this organization and the structure of the database, this resource is the second listed in the entire collection. This is an example of how the structure of a database may impact how certain resources are accessed.

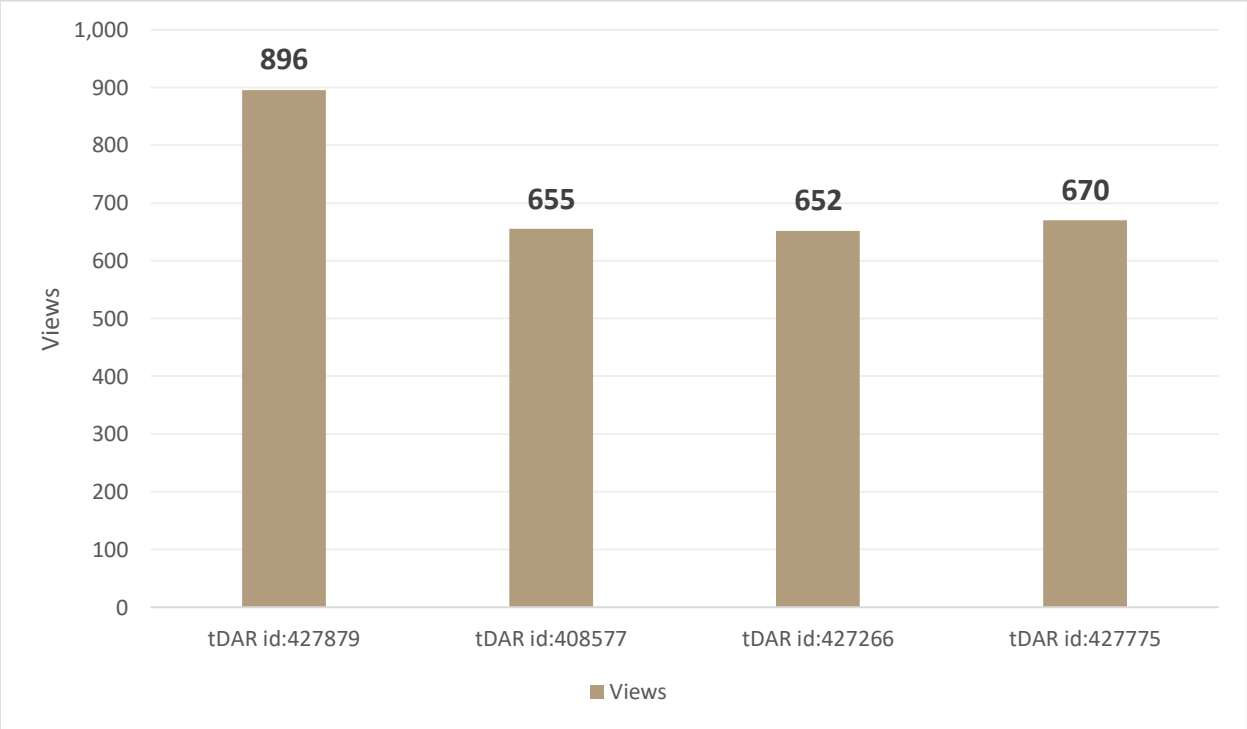

Figure 4.8. Total views of the SRP Digital Library sample resources as of Oct. 16, 2022.

Downloads per Resource Type

The sample that I examined showed varying rates of downloads for the files. Higher download rates for the confidential files suggest their use by CRM researchers, possibly branching into academic research as well. This is clear in Garcia's (personal communication

2022) comments about the frequency of requests from known CRM professionals and by tDAR users with academic institution email addresses (.edu). As shown in Figure 4.9, the confidential versions of the SRP Digital Library sample resources were downloaded with the same or greater frequency than the public, redacted versions. As mentioned above, the confidential files were uploaded first in 2017 while the redacted versions were not made available until approximately four years later. This surely impacts how these data are analyzed. As was the case with MimPIDD, downloads of the public versions of the resources could suggest both public engagement as well as access by professionals without the need for the redacted information. If anything, the fact that the redacted files were downloaded as many times as they were in only one year suggests that these files could soon become more downloaded than the confidential files, thereby demonstrating the value of having publicly accessible versions of the resources.

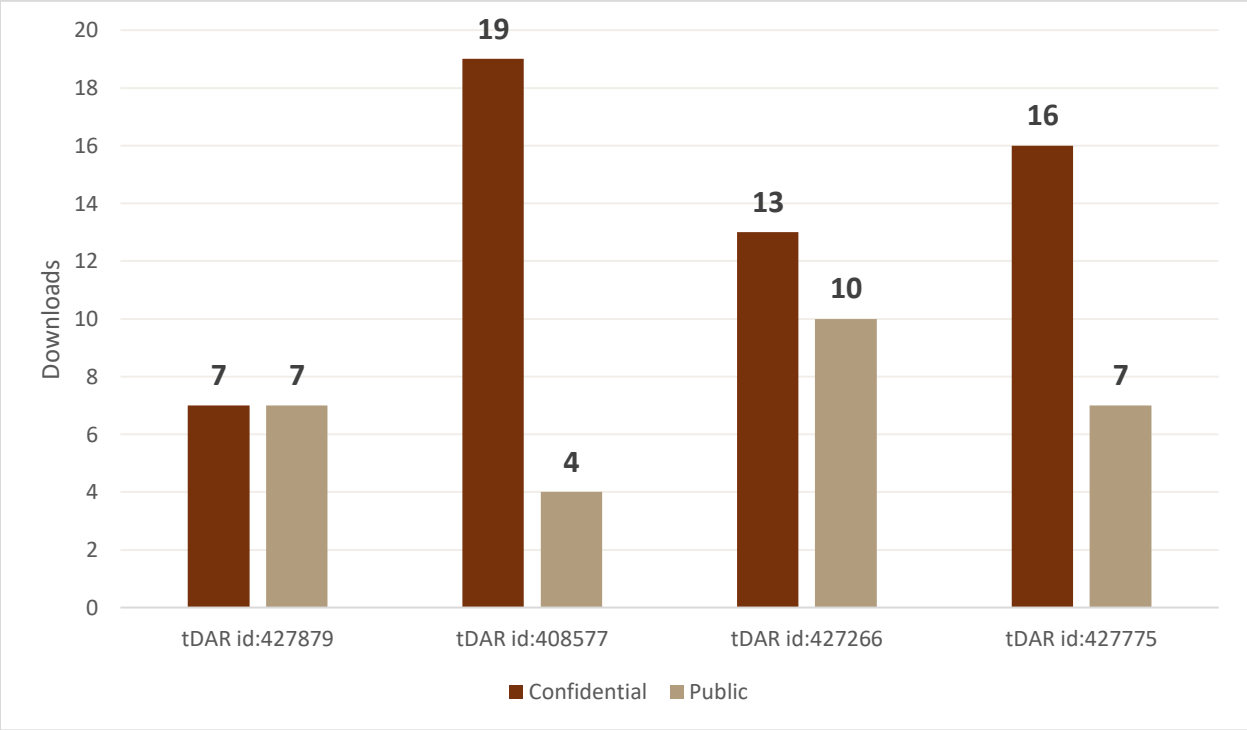

Figure 4.9. Total downloads of the confidential and public versions of the SRP Digital Library sample resources as of Oct. 16, 2022.

Reuse

Unlike the list curated by the MimPIDD administrators, there are no concrete data on reuse of the resources in the SRP Digital Library. Garcia (personal communication 2022) anecdotally shared that most requests to access the confidential documents in the SRP Digital Library come from CRM professionals, while some come from users with academic email addresses (.edu), although he does not always know for certain whether they are faculty or students.

Considering the scale of resources in the SRP Digital Library and the multitude of access requests, a usage survey sent to all individuals who downloaded materials could prove difficult. This sort of survey technique, however, could shed important light on who is using the SRP Digital Library and in what contexts. Since all tDAR users must provide an email address when registering their account and requesting access to confidential files, conducting such a survey would at least theoretically be possible, but still contingent on the users' responses. The SRP Digital Library administrator could collect these emails and send a survey with questions regarding what resources were accessed and for what reasons. A survey question specifically asking whether users downloaded and used the confidential or redacted version of the resources would be most helpful in identifying the impact of the access system. Possible questions could include:

- How many resources have you accessed?
- Which resources have you accessed?
- Did you download the confidential or redacted versions?

- In what context did you use the resource(s): Academic research, educational use, CRM, or general interest?
- If you downloaded the redacted version, did it include enough information for your purposes?
- Have you cited resources that you accessed through the SRP Digital Library? If so, please list the publications/uses.
- Do you know of any other resources that should be added to the SRP Digital Library?
- Are there any resources in the SRP Digital Library for which redacted versions should be created?

A more efficient way to collect reuse data would be to implement a pre-use survey in the request to access a resource specifically asking a user how and in what context they intend to reuse the data. This would facilitate data-gathering by making it a necessary facet of the access request process. Such pre-use surveys, however, would only be applicable to confidential resources as publicly available resources do not require the Request Access form and thus no means of requiring survey completion before downloading the resource. Data from pre-use surveys could then be compiled alongside post-use surveys to better understand exactly how data from the SRP Digital Library, and other databases that implement this reuse survey methodology, are being reused.

Chapter 5: Discussion and Conclusions

Summary of Analyses

Prior to the study presented in this thesis, there has been little investigation of the accessibility and reuse of archaeological databases housed in digital data management systems such as tDAR. The databases discussed in this thesis, MimPIDD and the SRP Digital Library, exemplify how goals for ethical data management and data reuse can be achieved simultaneously due to the databases' organized structures, consideration of content that may be sensitive, and data access models. Both MimPIDD and the SRP Digital Library employ ways to differentiate access to public and restricted versions of their resources; the effectiveness of this methodology is seen in the engagement metrics with both versions (Figures 4.1, 4.2, 4.8, and 4.9). Though the number of views and downloads for each resource does not directly indicate concrete reuse patterns, they represent the opportunity for data reuse (Huggett 2018:94). In the case of MimPIDD, survey results collected by the MimPIDD board show diverse instances of reuse using the data in their restricted-access research database. These instances of reuse range from public educational programs to academic publications by members of the MimPIDD board and other researchers from around the world (Figure 4.4). Furthermore, reuse of the SRP Digital Library appears to be predominantly by professional CRM archaeologists, but it extends into academic research. A survey similar to that employed by MimPIDD could provide quantitative and qualitative data necessary to discern specific patterns of reuse of the SRP Digital Library.

The compartmentalized and organized database structures, protection of culturally sensitive content, and differentiated access of both case studies can and should be applied to similar projects that aim to share archaeological data with professionals and the public alike. Designing dual-access databases like these requires time and dedication yet these should become

the norm for data sharing as digital data management systems will continue to play an increasing role in archaeological stewardship. Archaeologists must consistently remind themselves of the non-renewable nature of the archaeological record and take advantage of the digital tools at their disposal as the discipline continues to evolve. In the following sections I address the research questions posed in Chapter 1 in light of the results of my analyses of MimPIDD and the SRP Digital Library (Figure 5.1).

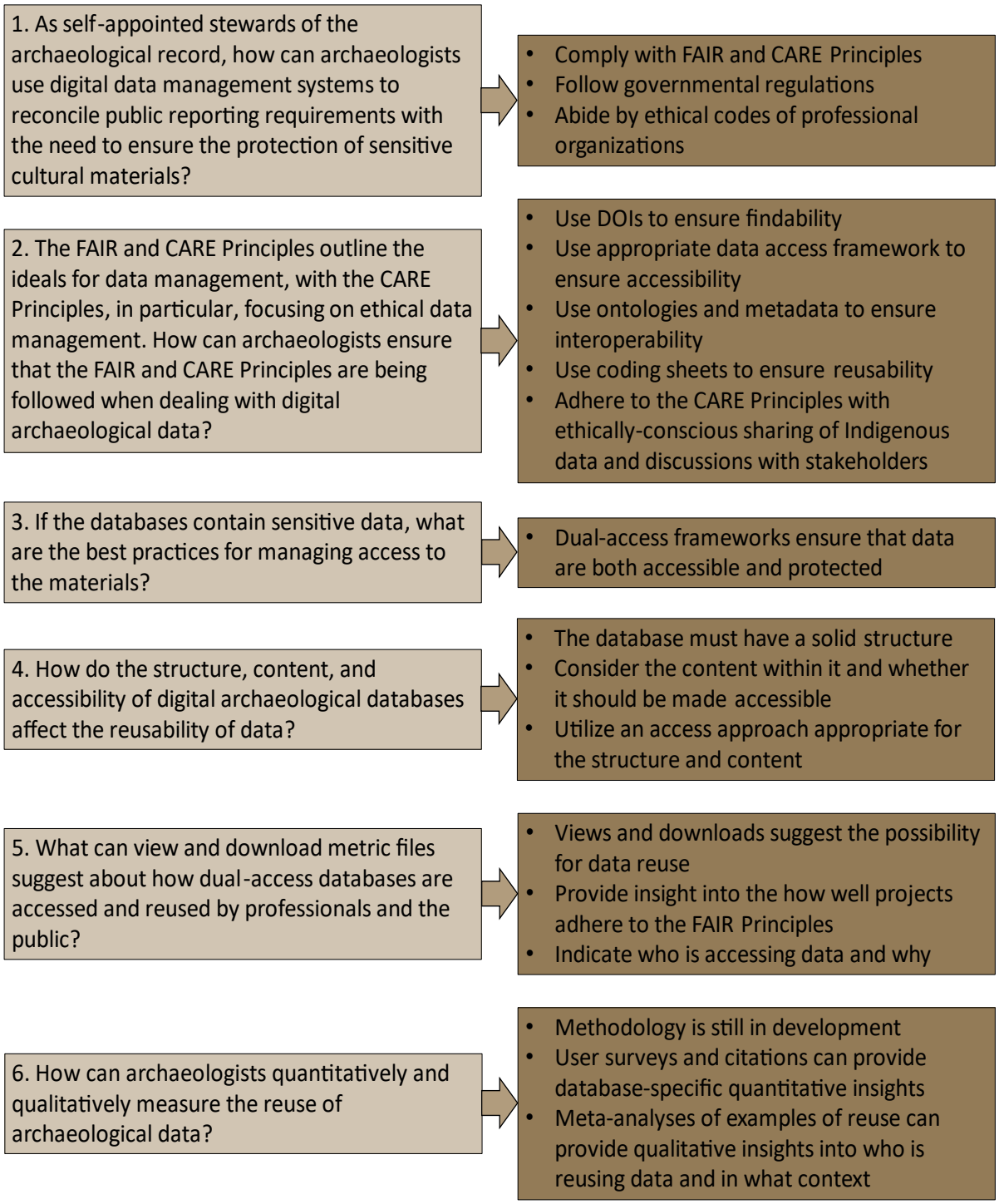

Figure 5.1. Research questions addressed in this thesis.

Ethical Data Management

As self-appointed stewards of the archaeological record, how can archaeologists use digital data management systems to reconcile public reporting requirements with the need to ensure the protection of sensitive cultural materials? This was the first research question posed in Chapter 1 and it is also the first question that archaeologists must confront before diving headfirst into digital archaeology. These goals are not mutually exclusive, but they are two sides of the same stewardship coin. In order to begin a discussion on stewardship with regard to digital data management, archaeologists must first come to terms with the amount of power they have in sharing information and the possible outcomes, positive and negative, that come with that responsibility. As stressed by the SAA, archaeologists have an important role to play as stewards who should strive to care and advocate for the irreplaceable archaeological record (Society for American Archaeology 2018).

The shift towards a digital archaeology both complicates and facilitates stewardship as the new information transfer universe poses unprecedented opportunities for data sharing alongside considerable technical and ethical challenges. The notion of accessibility, for one, is promising yet concerning. Sharing datasets, images, documents, and other resources through digital data management systems provides an unprecedented level of access to researchers, professionals, and the public. At the same time, this newfound openness creates additional issues ranging from the potential for information to get into the wrong hands much more easily to broader questions about who should have control over data access. MimPIDD and the SRP Digital Library both focus on this balance between accessibility and protection. The case studies exemplify how using a dual-access system can allow for public engagement and simultaneously protect sensitive data. By complying with the FAIR and CARE Principles, following

governmental regulations, and abiding by the ethical codes of professional organizations, archaeologists can find a balance in digital data management systems that were designed to address these considerations.

The FAIR and CARE Principles outline the ideals for data management, with the CARE Principles, in particular, focusing on ethical data management. How can archaeologists ensure that the FAIR and CARE Principles are being followed when dealing with digital archaeological data? Ethical issues are complex in the social sciences and humanities, especially archaeology given its focus on objects of cultural patrimony. This, paired with the fact that archaeology has historically prioritized data collection rather than consideration of the implications and aftereffects of curation and information management, means that archaeologists must consider the ethics of every decision they make before, during, and after they collect data. The use of digital data management systems makes it easier for researchers and CRM professionals to adhere to the technical and ethical standards for data management outlined by the FAIR and CARE Principles. tDAR, where MimPIDD and the SRP Digital Library are housed, uses DOIs and its own identifiers to ensure findability, an administrator-controlled data access framework (Figure 2.3) to ensure accessibility, ontologies and always-accessible resource metadata to ensure interoperability, and detailed metadata and coding sheets to ensure reusability. For these reasons, McManamon et al. (2018) explain that wider use of tDAR in CRM data curation would streamline Section 106 procedures and put to use the massive quantity of otherwise uninterpreted and disparate data. Both MimPIDD and the SRP Digital Library are ethically minded regarding sharing and protecting Indigenous data and emphasize reuse, factors that are in accordance with the CARE Principles.

As discussed in Chapter 2, the FAIR and CARE Principles are not exclusive to archaeology; they therefore lack the discipline-specific perspective necessary to stand alone. Moreover, while professional organizations like the SAA and AIA promote ethical principles, they lack the nuance found in digital data management. As recommended by Nicholson et al. (2021) and Richardson (2018:70), professional institutions like the SAA should develop and promote measurable and enforceable standards for archaeology-specific ethical digital data management. This is something that must be promptly addressed as this new methodology is gaining traction. In the meantime, archaeologists should continue to engage in discussions about how the FAIR and CARE Principles relate to their data and hold each other accountable for ensuring the principles are being implemented and enforced.

If the databases contain sensitive data, what are the best practices for managing access to the materials? As exemplified by MimPIDD and the SRP Digital Library, a dual access approach ensures that data are both accessible and protected by providing two versions of the data: one that is fully accessible yet partially redacted, and a second that contains complete, uncensored data but is only accessible via a request to the database administrator. The two case studies were chosen specifically to show that different methodologies are necessary and effective for different resource types and audiences. Archaeologists must consider this when designing their own databases and access models. Regardless, databases must adhere to the FAIR and CARE Principles along with any discipline-specific standards established in the future.

The tDAR access framework (Figure 2.3) allows administrators to control access, but each case study employs a slightly different methodology for ensuring that data are only shared with those qualified to access the information. MimPIDD's double-database approach requires a written request to the administrator to access the full research database wherein the user must

acknowledge ethical issues associated with the data themselves alongside broader ethical principles about data sharing. The SRP Digital Library, on the other hand, uses a file-by-file request-to-access system wherein requests are vetted by the SRP archaeologist who weighs the prospective research angle against the sensitive locational and burial content in the resources. Additionally, some resources in the collection have redacted versions that can be downloaded by any tDAR user without the need to request access. Access to restricted resources could be improved with a more seamless request form. Such a form could contain all fields relevant to the specific database and the intended goals for reuse rather than the open-ended text box currently used in tDAR's access request form (Figure 4.7). This would ensure that requests contain all information necessary for the administrator to decide whether to grant access and allow requests to be processed more efficiently. Additionally, the form could include questions regarding how and in what context the user intends to reuse the data as well as user information such as their name, institution, and email address, which would make it easier for database administrators to reach back out to users regarding their reuse of data.

Data Reuse

How do the structure, content, and accessibility of digital archaeological databases affect the reusability of data? Each of these characteristics, which are described in detail in Chapter 3, builds upon the one that precedes it, creating a pyramid-like hierarchy (Figure 5.2). In order for a database to effectively foster data reuse, it must have a solid structure on a digital data management platform such as tDAR to handle content, consider the content within it and whether it should be made accessible, and utilize an access approach appropriate for the structure

and content. These characteristics function alongside the FAIR and CARE Principles to facilitate reuse.

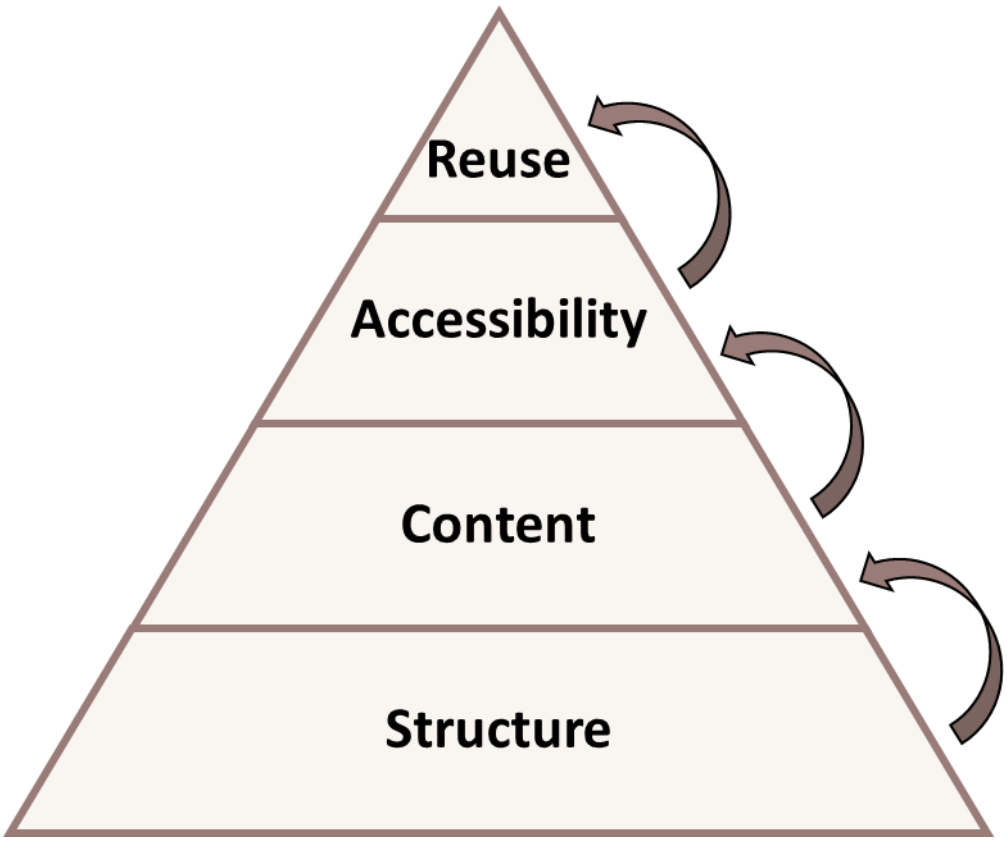

Figure 5.2. Hierarchy of database structure, content, and accessibility in fostering reuse.

MimPIDD was designed to facilitate large-scale analyses of Mimbres pottery and allow for public engagement with the resources within it, so it follows that its structure consists of two databases: a research version and a public version. This dual-database structure allows the administrators to organize the content within MimPIDD, with the public version containing 2,188 resources and the research version having an additional 7,395 images from private collections as well as detailed locational and burial-related attribute data. MimPIDD's accessibility is based on its content, with the more sensitive data of the research version being

restricted to only those whose access requests were approved by the MimPIDD administrator. This differentiated accessibility allows for reuse by providing many images, basic attribute data, and educational materials designed for public engagement and by ensuring that researchers can access data that are otherwise inaccessible due to the desire to protect sensitive data and permissions by private collectors.

The SRP Digital Library had similar motives in its development, with the desire to create a collection of resources related to Hohokam archaeology that can be accessed and reused by academics, CRM professionals, and the public. The collection contains 776 resources, many of which have confidential and redacted versions that black out sensitive locational information and images. The accessibility of the SRP Digital Library is based on this system of differentiated versions of resources; thus, users must request access to the confidential files by submitting a request to access form that must be approved by the SRP Digital Library administrator. Like MimPIDD, differentiating access in this way allows for the public to engage with the data without sharing too much and putting sites, artifacts, or descendant groups at risk while also providing researchers and CRM professionals with more detailed data.

What can view and download metric files suggest about how dual-access databases are accessed and reused by professionals and the public? The analyses presented in this thesis accord with Huggett's (2018:94) argument that view and download metrics only suggest the possibility of data reuse. That being said, view and download metrics can still provide meaningful insight into how well projects adhere to the FAIR Principles. These data are useful to the administrators of particular projects and to the digital data management platforms themselves as they indicate the frequency of engagement. As outlined in Chapter 3, we might expect that views and downloads of restricted-access resources reflect professional and academic

engagement with the materials while views and downloads for fully accessible resources indicate more casual research interests and public engagement. These assumptions were supported by the reuse data provided by the MimPIDD board. The interaction statistics of both MimPIDD and the SRP Digital Library are consistent with the idea that differentiating access can lead to engagement with both the restricted and public versions of the resources (Figures 4.1, 4.2, 4.8, and 4.9). While less straightforward than uploading and managing only one form of the same data, the statistics demonstrate that it is worth differentiating access.

How can archaeologists quantitatively and qualitatively measure the reuse of archaeological data? As discussed throughout this thesis, the current methodology for measuring reuse remains underdeveloped and insufficient for comprehensive analyses. User surveys, such as those implemented by the MimPIDD board and suggested for the SRP Digital Library in Chapter 4, can ask certain questions to provide administrators with database-specific insights. Surveys can be powerful tools, yet they may not be easy to implement, and their feasibility is predicated on the ability to reach out to the relevant users. Moreover, it is unlikely that every recipient of a survey will complete it, so valuable data might not be recorded. Thus, pre-use surveys implemented at the access request stage for confidential resources could elucidate patterns of prospective reuse that could be confirmed with post-use surveys. This survey methodology, however, notably leaves out responses from users downloading public resources that do not require permission to download. Citations and DOIs in published writing can indicate instances of reuse to database administrators, yet such references are unfortunately not the current norm in archaeological research, with no standardized requirements for doing so (Huggett 2018:95). Besides, these sorts of references would only identify *published* instances of reuse and thus leave out examples of reuse outside of professional archaeology. Many students

and avocational researchers find ways to use these data, yet these uses are undiscovered and undiscoverable to date.

The list of reuses provided by the MimPIDD board shows contextually and temporally widespread instances of reuse by individuals with varying connections to the collection. Among the 53 downloads of the full research dataset (Figure 4.1), there were 26 unique first authors identified in the reuse list. If we assume that each of these 26 authors downloaded the full research dataset, they represent almost exactly half of the total downloads. This finding would need to be substantiated by further analyses of other resources in this collection before any definitive pattern of reuse might be discerned, but it suggests a substantial ratio of downloads-to-reuses.

Currently there are no concrete reuse data available for the SRP Digital Library, and such data could prove difficult to obtain given the nature of the resources within the database. Unlike MimPIDD, which contains raw datasets and images that can be reused, the SRP Digital Library functions as its name suggests: it is a library with resources by various authors from various sources. Where MimPIDD's raw data can be cited, citing the resources in the SRP Digital Library does not necessitate acknowledging the database; rather, documents themselves can be cited. This rules out measuring reuse through citations and is precisely why a reuse survey, including the questions in Chapter 4, could prove beneficial. Additionally, at the time of writing there is no system in place to track access requests within tDAR, so if administrators wish to track such data, they must do so themselves. Developing a tool to track this data would provide beneficial information on the frequency of requests to access resources, and thus supplement view and download metrics in understanding the potential for data reuse.

Tracking instances of reuse across all contexts through pre- and post-use surveys as well as tracking DOIs and other citations that connect published instances of reuse to the original data would provide the clearest picture of data reuse and is especially relevant to projects such as MimPIDD, which focus so heavily on public engagement with their materials. Such statistics could be used to secure grant funding and improve the user experience of these databases. Qualitative analyses of instances of reuse are similarly difficult to perform yet they can demonstrate the FAIRness and overall effectiveness of a database. These analyses can examine who reused data, the context of reuse (academic research, CRM work, public education, etc.), how data were used, and how the structure, content, and accessibility of a database facilitated reuse. The qualitative analysis of MimPIDD reuse (discussed in Chapter 4) showed that data were used in a variety of contexts, though mostly in the form of academic research, by a variety of individuals around the world. Understanding instances of reuse can highlight the positive characteristics of the databases as well as indicate potential areas for improvement.

Significance of Research

This thesis presents data demonstrating that the structure, content, and accessibility of MimPIDD and the SRP Digital Library create opportunities for meaningful and necessary archaeological data reuse. These projects should serve as blueprints for the future of ethical and accessible data management in archaeological research as we consider our role as stewards and our impact on the archaeological record. The data on reuse available at the time of writing indicate the importance of following the FAIR and CARE Principles and using dual-access models in facilitating archaeological data reuse. Further development of a survey-based methodology for measuring reuse is the next step in the expanded use of digital data

management systems. Such a methodology would be greatly beneficial for the platforms themselves as well as other scientists interested in measuring qualitative data reuse.

As our world continues to move into the digital realm, it follows that archaeology should do the same. This thesis demonstrates that digital data management systems like tDAR allow archaeologists to preserve, share, and reuse information in new ways. With the application of thoughtful ethical principles as exemplified in these case studies, archaeology can become more open and accessible to academics and the public. Archaeology has always been considered a dirt-based and hands-on science, but digital data reuse may be changing that reputation forever.

Future Research

The digital sphere is constantly changing as new methods for data collection, publication, and access evolve. Future investigations into similar topics will inevitably continue to push the paradigm shift forward. As suggested in Chapter 4, a survey to collect quantitative and qualitative data on the reuse of the materials in the SRP Digital Library could be implemented to elucidate patterns of reuse and support the efficacy of their data accessibility model. Moreover, a larger-scale survey of MimPIDD reuse could include reaching out not only to users who were given access to the full research database but also individuals who accessed the public database, which could further highlight the widespread use of this frequently accessed collection. Surveys like these are currently the easiest and most accessible way of measuring reuse, but is there a better way? Further investigations into the development of a methodology for consistently and accurately tracking data reuse from digital data management platforms are the necessary next step.

This thesis focused specifically on two case studies from the same platform, tDAR. Future research could use similar methods to analyze and interpret the structure, content, accessibility, and reuse of digital databases housed in other platforms like Open Context and the ADS. While investigations of Open Context resources could be limited by their lack of view and download metrics due to their privacy policy, the unique style of these resources as peer-reviewed published data could highlight the pros and cons of each mode of data sharing. Does the way that data are made available for reuse impact how they are accessed and reused?

Finally, another necessary next step regarding digital archaeological data access and reuse is the development of discipline-specific digital data management standards, specifically relating to citing the original data and ethical considerations associated with cultural objects. This thesis included discussion and citation of many digital resources, yet a lack of thorough digital data citation standards made it difficult to ensure that all resources are findable. As suggested by a number of scholars (e.g., Altschul and Klein 2022; Nicholson et al. 2021; Richardson 2018:70), the SAA and other professional and governmental organizations should establish measurable and enforceable standards to supplement the FAIR and CARE Principles. Discussions about these standards should involve a variety of stakeholder voices, including academic researchers, educators, CRM professionals, State Historic Preservation Officers (SHPOs), and most importantly, Indigenous communities and Tribal Historic Preservation Officers (THPOs). An ethically conscious and reuse-focused archaeology lies just beyond the horizon, but it can only be reached through collaboration.

References Cited

Advisory Council on Historic Preservation
2018 Section 106 Archaeology Guidance. Electronic document, https://www.achp.gov/Protecting-Historic-Properties/Section_106_Archaeology_Guidance, accessed October 6, 2022.

Altschul, Jeffrey H., and Terry H. Klein
2022 Forecast for the US CRM Industry and Job Market, 2022-2031. *Advances in Archaeological Practice*:1-16.

Archaeology Southwest
2022 What We Do, *Archaeology Southwest*. Electronic document, https://www.archaeologysouthwest.org/about/what-we-do/, accessed September 29, 2022.

2017 Hohokam or Huhugam?, *Archaeology Southwest*. Electronic document, https://www.archaeologysouthwest.org/exhibit/online-exhibits/pieces-puzzle/piece-1/#:~:text=Huhugam%20is%20an%20O'odham,ancient%20groups%20in%20the%20Southwest, accessed September 27, 2022.

Atici, Levent, Sarah Whitcher Kansa, Justin Lev-Tov, and Eric C Kansa
2012 Other People's Data: A Demonstration of the Imperative of Publishing Primary Data. *Journal of Archaeological Method and Theory* 20(4):663-681.

Beck, Anthony, and Cameron Neylon
2012 A vision for Open Archaeology. *World Archaeology* 44(4):479-497.

Bevan, Andrew
2015 The Data Deluge. *Antiquity* 89(348):1473-1484.

Bollwerk, Elizabeth
2015 Co-Creation's Role in Digital Public Archaeology. *The Society for American Archaeology* 3(3):223-234.

Bonacchi, Chiara, and Gabriel Moshenska
2015 Critical Reflections on Digital Public Archaeology. *Internet Archaeology*.

Bria, Rebecca E., and Erick Casanva Vasquez
2022 Digital Archaeology and Storytelling as a Toolkit for Community-Engaged Archaeology. In *Critical Archaeology in the Digital Age: Proceedings of the 12th IEMA Visiting Scholar's Conference*, edited by Kevin Garstki, pp. 49-65. The Cotsen Institute of Archaeology Press, Los Angeles.

Brin, Adam

2015 Taking a look back at 2014, tDAR News. Electronic document, https://www.tdar.org/news/2015/01/taking-a-look-back-at-2014/, accessed October 16, 2022.

2016 Taking a look back at 2015, tDAR News. Electronic document, https://www.tdar.org/news/2016/01/taking-a-look-back-at-2015/, accessed October 16, 2022.

2017 Taking a look back at 2016, tDAR News. Electronic document, https://www.tdar.org/news/2017/01/taking-a-look-back-at-2016/, accessed October 16, 2022.

2018 Looking Back at 2017, Part II: tDAR, Content, and Usage, tDAR News. Electronic document, https://www.tdar.org/news/2018/01/looking-back-at-2017-part-ii-tdar-content-and-usage/, accessed October 16, 2022.

Brody, J. J.
2004 *Mimbres Painted Pottery*. Revised edition. ed. Santa Fe: School of American Research Press, Santa Fe, New Mexico.

Cambridge Core
2020 Download metrics of "The Social Significance of Mimbres Painted Pottery in the U.S. Southwest" by Hegmon, Michelle, Will G. Russell, Kendall Baller, Matthew A. Peeples, and Sarah Striker. Electronic document, https://www.cambridge.org/core/journals/american-antiquity/article/abs/social-significance-of-mimbres-painted-pottery-in-the-us-southwest/E819E6098CA2C969EC0359E7D33D3CBF#metrics, accessed September 27, 2022.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, and Maui Hudson
2020 The CARE Principles for Indigenous Data Governance. *Data Science Journal* 19(1):1-12.

Carroll, Stephanie Russo, Edit Herczog, Maui Hudson, Keith Russell, and Shelley Stall
2021 Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Scientific Data* 8(1):108.

Cordell, Linda S., and Maxine E. McBrinn
2016 *Archaeology of the American Southwest*. 3 ed. Routledge, New York.

Cherry, John F.
2011 Still not digging, much. *Archaeological Dialogues* 18(1):10-17.

Clarke, Mary

2015 The Digital Dilemma: Preservation and the Digital Archaeological Record. *Advances in Archaeological Practice* 3(4):313-330.

Currie, Adrian
2018 *Rock, Bone, and Ruin: An Optimist's Guide to the Historical Sciences / Adrian Currie*. Life and mind: philosophical issues in biology and psychology. The MIT Press, Cambridge, Massachusetts.

Dennis, L. Meghan
2020 Digital Archaeological Ethics: Successes and Failures in Disciplinary Attention. *Journal of Computer Applications in Archaeology* 3:210-218.

Digital Antiquity
2011 Access Rights in tDAR, the Digital Archaeological Record (tDAR) Confluence. Electronic document, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557185/Access+Rights+in+tDAR, accessed October 16, 2022.

2013 Managing Security and Access, the Digital Archaeological Record (tDAR) Confluence. Electronic document, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557219/Managing+Security+and+Access, accessed September 25, 2022.

2018a Compliance, the Digital Archaeological Record (tDAR). Electronic document, https://www.tdar.org/why-tdar/compliance/, accessed August 21, 2022.

2018b History, the Digital Archaeological Record (tDAR). Electronic document, https://www.tdar.org, accessed October 6, 2022.

2020 Ontology, the Digital Archaeological Record (tDAR). Electronic document, https://www.tdar.org/using-tdar/creating-and-editing-resources/ontology-create-and-edit/, accessed December 9, 2022.

2021 Upload Toolkit, the Digital Archaeological Record (tDAR). Electronic document, https://www.tdar.org/using-tdar/upload-toolkit/#tDarResourceTypes, accessed September 15, 2022.

Fagan, Brian
1995 Timelines: Archaeology's Dirty Secret. *Archaeology* 48(4):14-17.

Faniel, Ixchel M., Anne Austin, Eric Kansa, Sarah Whitcher Kansa, Phoebe France, Jennifer Jacobs, Ran Boytner, and Elizabeth Yakel
2018 Beyond the Archive: Bridging Data Creation and Reuse in Archaeology. *Advances in Archaeological Practice* 6(2):105-116.

Finegold, Andrew

2019 Objects without Texts: Mimbres Painted Bowls and the Problematics of Interpretation. *Art History* 42(2):216-241.

Fish, Suzanne K., and Paul R. Fish
2008 The Hohokam Millennium. In *The Hohokam Millennium*, edited by Suzanne K. Fish, and Paul R. Fish, pp. 1-11. School for Advanced Research Press, Santa Fe, New Mexico.

Garstki, Kevin
2020 *Digital Innovations in European Archaeology*. Cambridge University Press, Cambridge, United Kingdom.

2022 Introduction: Challenges of a Critical Archaeology in the Modern World. In *Critical Archaeology in the Digital Age: Proceedings of the 12th IEMA Visiting Scholar's Conference*, edited by Kevin Garstki, pp. 1-9. The Cotsen Institute of Archaeology Press, Los Angeles.

Gattiglia, Gabriele
2015 Think big about data: Archaeology and the Big Data challenge. *Archäologische Informationen* 38:113-124.

Geser, Guntram, Julian D. Richards, Flavia Massara, and Holly Wright
2022 Data Management Policies and Practices of Digital Archaeological Repositories. *Internet Archaeology* 59.

Hegmon, Michelle, James R. McGrath, and Marit K. Munson
2017 The Potential and Pitfalls of Large Multi-Source Collections: Insights from the Analysis of Mimbres Gender Imagery. *Advances in Archaeological Practice* 5(2):138-146.

Hegmon, Michelle, Will G. Russell, Kendall Baller, Matthew A. Peeples, and Sarah Striker
2021 The Social Significance of Mimbres Painted Pottery in the U.S. Southwest. *American Antiquity* 86(1):23-42.

Hill, J. Brett, Patrick D. Lyons, Jeffery J. Clark, and William H. Doelle
2015 The "Collapse" of Cooperative Hohokam Irrigation in the Lower Salt River Valley. *Journal of the Southwest* 57(4):609-674.

Hill, J. Brett
2019 *From Huhugam to Hohokam: Heritage and Archaeology in the American Southwest*. Lexington Books, Lanham, Maryland.

Huggett, Jeremy
2018 Reuse Remix Recycle: Repurposing Archaeological Digital Data. *Advances in Archaeological Practice* 6(2):93-104.

2020 Is Big Digital Data Different? Towards a New Archaeological Paradigm. *Journal of Field Archaeology* 45:S8-S17.

2022 Is Less More? Slow Data and Datafication in Archaeology. In *Critical Archaeology in the Digital Age: Proceedings of the 12th IEMA Visiting Scholar's Conference*, edited by Kevin Garstki, pp. 97-110. The Cotsen Institute of Archaeology Press.

Isaksen, Leif, Martinez, Kirk, Gibbins, Nicholas, Earl, Graeme and Keay, Simon
2009 Linking archaeological data. Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference:1-18.

Kansa, Eric C., and Sarah Whitcher Kansa
2013 We All Know That a 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication. *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 1(1):88-97.

Kansa, Eric C.
2016 Click Here to Save the Past. In *Mobilizing the Past for a Digital Future*, edited by Erin Walcek Averett, Jody Michael Gordon, and Derek B. Counts, pp. 443-472. The Digital Press at the University of North Dakota, Grand Forks.

Kersel, Morag M.
2015 An Issue of Ethics? Curations and the Obligations of Archaeology. *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 3(no. 1):77-79.

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

Koster, Lukas, and Saskia Woutersen-Windhouwer
2018 FAIR Principles for Library, Archive and Museum Collections: A proposal for standards for reusable collections. *Code4Lib Journal* 40.

Marquardt, William H., Anta Montet-White, and Sandra C. Scholtz
1982 Resolving the Crisis in Archaeological Collections Curation. *American Antiquity* 47(no. 2):409-418.

McManamon, Francis
2018 Making Data Easier to Discover, Access, and Use as Part of Section 106 Streamlining. *tDAR News*, October 8. https://www.tdar.org/news/2018/10/making-data-easier-to-discover-access-and-use-as-part-of-section-106-streamlining/, accessed December 11, 2022.

McManamon, Francis P., Keith W. Kintigh, Leigh Anne Ellison, and Adam Brin

2017 tDAR: A Cultural Heritage Archive for Twenty-First-Century Public Outreach, Research, and Resource Management. *Advances in Archaeological Practice* 5(3):238-249.

Nelson, Margaret C., and Patricia A. Gilman
2017 Mimbres Archaeology. In *The Oxford Handbook of Southwest Archaeology*, edited by Barbara Mills, and Severin Fowles, pp. 265-283. Oxford University Press, Oxford, United Kingdom.

Nicholas, George, and Julie Hollowell
2007 Ethical Challenges to a Postcolonial Archaeology: The Legacy of Scientific Colonialism. In *Archaeology and Capitalism: From Ethics to Politics*, edited by Yannis Hamilakis, and Phillip Duke, pp. 59-82. Routledge, New York.

Nicholson, Christopher M., Rachel Fernandez, and Jessica Irwin
2021 Digital Archaeological Data in the Wild West: the challenge of practising responsible digital data archiving and access in the United States. *Internet Archaeology* 58. https://doi.org/10.11141/ia.58.22.

Perry, Sara
2018 Why Are Heritage Interpreters Voiceless at the Trowel's Edge? A Plea for Rewriting the Archaeological Workflow. *Advances in Archaeological Practice* 6(3):212-227.

Piwowar, Heather A., and Todd J. Vision
2013 Data reuse and the open data citation advantage. *PeerJ* 1:e175. https://doi.org/10.7717/peerj.175.

Richards, Julian D.
2002 Digital Preservation and Access. *European Journal of Archaeology* 5:343-366.

2015 Ahead of the curve: adventures in e-publishing in *Internet Archaeology*. *Archäologische Informationen* 38:63-71.

Richards, Julian D., Ulf Jakobsson, David Novák, Benjamin Štular, and Holly Wright
2021 Digital Archiving in Archaeology: The State of the Art. *Internet Archaeology* 58.

Richardson, Lorna-Jane
2018 Ethical Challenges in Digital Public Archaeology. *Journal of Computer Applications in Archaeology* 1(1):64-73.

Roosevelt, Christopher H, Peter Cobb, Emanuel Moss, Brandon R Olson and Sinan Ünlüsoy
2015 Excavation Is ~~Destruction~~ Digitization: Advances in Archaeological Practice. *Journal of Field Archaeology* 40:325-346.

Salt River Project

n.d. A history of canals in Arizona. Electronic document, https://www.srpnet.com/about/history/canal-history, accessed October 2, 2022.

Shanks, Michael, and Christopher Tilley
1992 Re-Constructing Archaeology: Theory and Practice. 2nd ed. Routledge, London.

Shepard, Lindsay
2015 Using a Multidisciplinary Approach to Interpret Artifacts. Blog entry, *Archaeology Southwest*, Tucson, Arizona.

Society for American Archaeologists
2018 Principles of Archaeological Ethics. Electronic document, https://www.saa.org/career-practice/ethics-in-professional-archaeology, accessed October 6, 2022.

Supernant, Kisha, and Gary Warrick
2014 Challenges to Critical Community-based Archaeological Practice in Canada. *Canadian Journal of Archaeology / Journal Canadien d'Archéologie* 38(2):563-591.

tDAR (the Digital Archaeological Record)
22070 Mimbres Pottery Database (Public) (n.d.). Electronic database, the Digital Archaeological Record (tDAR), https://core.tdar.org/collection/22070/mimbres-pottery-images-digital-database-with-search, accessed June 30, 2022.

30428 Digital Archive of Huhugam Archaeology (n.d.). Electronic database, the Digital Archaeological Record (tDAR), https://core.tdar.org/collection/30428/digital-archive-of-huhugam-archaeology-daha, accessed October 2, 2022.

57630 Salt River Project Digital Library (n.d.). Electronic database, the Digital Archaeological Record (tDAR), https://core.tdar.org/collection/57630/salt-river-project-digital-library, accessed October 16, 2022.

377853 Hegmon, Michelle and Will Russell. A Quantitative Method for Recognizing the Same Hands in Mimbres Black-on-White Pottery (2013). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/377853/a-quantitative-method-for-recognizing-the-same-hands-in-mimbres-black-on-white-pottery, accessed September 30, 2022.

381421 Hegmon, Michelle. Ethics and Permission to Access MimPIDD (2022). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/381421/ethics-and-permission-to-access-mimpidd, accessed September 29, 2022.

381505 MimPIDD Research Database for Download (2019). Electronic database, the Digital Archaeological Record (tDAR), https://core.tdar.org/dataset/381505/mimpidd-research-database-for-download, accessed October 1, 2022.

408577 Clark, Tiffany C., and Kathleen Henderson. Cultural Resources Survey for the SRP Palo Verde to Pinal West 500 kV Transmission Lines, Maricopa and Pinal Counties, Arizona (2007). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/408577/cultural-resources-survey-for-the-srp-palo-verde-to-pinal-west-500-kv-transmission-lines-maricopa-and-pinal-counties-arizona, accessed September 28, 2022.

427266 Wilcox, David R., Jerry B. Howard, and Rueben H. Nelson. One Hundred Years of Archaeology at La Ciudad de Los Hornos (1990). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/427266/one-hundred-years-of-archaeology-at-la-ciudad-de-los-hornos, accessed September 28, 2022.

427775 Howell, Todd L. Archaeological Investigations at Los Guanacos: Exploring Cultural Changes in Late Hohokam Society (1993). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/427775/archaeological-investigations-at-los-guanacos-exploring-cultural-changes-in-late-hohokam-society, accessed September 28, 2022.

427879 Clark, Jeffery J., and James M. Vint. 2000 Years of Settlement in the Tonto Basin: Overview and Synthesis of the Tonto Creek Archaeological Project (2004). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/427879/2000-years-of-settlement-in-the-tonto-basin-overview-and-synthesis-of-the-tonto-creek-archaeological-project, accessed September 28, 2022.

455439 Kamp-Whittaker, April and Michelle Hegmon. K-3 Education Resources – Mimbres Pottery Digital Database (2020). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/project/455439/k-3-education-resources-mimbres-pottery-digital-database, accessed August 10, 2022.

455444 Kamp-Whittaker, April and Michelle Hegmon. College/University Education Resources – Mimbres Pottery Digital Database (2020). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/project/455444/collegeuniversity-education-resources-mimbres-pottery-digital-database, accessed August 10, 2022.

455448 Kamp-Whittaker, April and Michelle Hegmon. Grades 4-12 Education Resources – Mimbres Pottery Digital Database (2020). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/project/455448/grades-4-12-education-resources-mimbres-pottery-digital-database, accessed August 10, 2022.

455455 Hegmon, Michelle. The Social Significance of Mimbres Painted Pottery in the U.S. Southwest (2021). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/project/455455/the-social-significance-of-mimbres-painted-pottery-in-the-us-southwest, accessed September 27, 2022.

455490 Hegmon, Michelle. MIMPIDD Users' Guide (2020). Electronic document, the Digital Archaeological Record (tDAR), https://core.tdar.org/document/455490/mimpidd-users-guide, accessed January 2, 2022.

Trigger, Bruce G.
1984 Alternative Archaeologies: Nationalist, Colonialist, Imperialist. *Man* 19(3):355-370.

VanValkenburgh, Parker, and J. Andrew Dufton
2020 Big Archaeology: Horizons and Blindspots. *Journal of Field Archaeology* 45:S1-S7.

Wallrodt, John
2016 Why Paperless: Technology and Changes in Archaeological Practice, 1996-2016. In *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, edited by Erin Walcek Averett, Jody Michael Gordon, and Derek B. Counts, pp. 33-50. The Digital Press at the University of North Dakota, Grand Forks.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J. G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A. C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons
2016 The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3(1):160018.

Wylie, Alison
2017 How Archaeological Evidence Bites Back: Strategies for Putting Old Data to Work in New Ways. *Science, Technology, & Human Values* 42(2):203-225.

York, J., M. Gutmann, and F. Berman
2018 What Do We Know About the Stewardship Gap. *Data Sci J* 17.