Anna C. Groves. A Model for Collaborative Relationships In The Digital Humanities. A Master's Paper for the M.S. in I.S degree. December, 2019. 76 pages. Advisor: Brian Sturm

This paper presents a Digital Humanities Collaborative Research Model, describing key linkages between four communities in digital humanities (DH): humanities scholars, data scientists, library/information services, and administrators. The model arises from literature review and original research. Original data was collected via interviews with seven humanities scholars from five institutions. Each participant is a well-known scholar experienced in DH research methods. Interviews evaluate challenges faced by DH scholars in these areas: data access, computational methods, library support services, publication incentives, and collaboration between scholars and other experts. Interview data was analyzed using grounded theory. Key findings: (1) DH scholarship depends on expert human collaboration within the DH ecosystem, especially between scholars and data scientists, as well as scholars and library/information specialists; (2) computational tools enhance, rather than substitute for, human collaboration; (3) publications incentives should be shared among the three major players in DH: humanities experts, data science experts, and information seeking experts.

Headings:

Communities of practice

Digital library administration

Digital library use studies

Electronic information resource literacy

Humanities libraries

Information economy

A MODEL FOR COLLABORATIVE RELATIONSHIPS IN THE DIGITAL HUMANITIES

by
Anna C. Groves

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

December 2019

Approved by

_____

Brian Sturm

Table of Contents

Introduction

The adoption of digital tools into the scholarly research process is an ongoing trend, generating changes in scholarly methods across university departments and disciplines (Unsworth et al., 2006). However, this evolutionary process and its effects have not been uniform across disciplines, departments, and institutions. In particular, the humanities have struggled to define the role of technology within their domains. Over the last few decades, the term "Digital Humanities" (DH) has emerged to describe computationally-based methods and projects within the humanities. The term covers an enormously broad range of activities, epistemologies, and organizations, and new articles are published on a regular basis attempting to describe and define Digital Humanities. Unfortunately, there is no real consensus on a unified and coherent model of the domain (Alvarado, 2011). The goals and priorities seem to be as idiosyncratic as the institutions and scholars involved. Defining a stable common denominator between project types, sets of digital tools, and data sources is difficult.

The ambiguity of DH presents a serious dilemma for institutions seeking to understand how to encourage and invest in DH scholarship and design supportive services for scholars interested in participating. A great deal of serious research has already been done to explore various approaches to the dilemma, examining everything from methodologies, ontologies, infrastructures, technologies, and support networks. This paper cannot hope to rival the full spectrum of these previous efforts. Rather, the goal is

to present a new model of analysis for the DH domain and an up-to-date snapshot of the barriers humanities scholars face in their use of computationally-based processes and methods. The basic research question is: From the perspective of scholars in the humanities, what are the primary barriers to the adoption and growth of digital methods and tools? It is hoped that the exploration of this question will provide a positive contribution to the decision-making process of university administrations, library boards, funding organizations, and other key institutions with a stake in the outcome of DH scholarship.

This paper presents a model for analyzing the relationships between four communities of practice involved in DH: humanities scholars, data scientists, library and information service professionals, and administrators overseeing policy and funding. It focuses on issues between the four domains from the perspective of digital humanities scholars. The issues of DH participants in the other three domains are presented from the perspective of humanities scholars, and in the context of their relationships to those scholars. However, the analytic model presented is intended to enable future research to complete evaluation from each perspective. With a better understanding of the relationships between the four primary communities of practice in DH, it is hoped that a sustainable method of exchange and collaboration is achievable.

The paper begins with a literature review section to lay the foundation for the analytic model and to provide an overview of issues unique to each of the four domains within it. A data analysis section follows, presenting original research derived from interviews with humanities scholars. These semi-structured interviews were conducted with seven humanities scholars from five major academic or research institutions in the humanities

field. Each participant is a well-known humanities scholar, familiar with current DH research methods and access to a major library with an extensive digital humanities collection. The interviews were designed to analyze the experiences and motivations of these scholars in conducting DH research projects. The structure of the interview questions was informed by extensive background reading outlined in the literature review section of this paper. The interviews covered topics including computational research methods employed, barriers and challenges encountered, library support services used or recommended, collaboration challenges between humanities scholars and computational/archive/library experts, and incentives tied to the digital publication process. The interviews were transcribed, with the content analyzed using a grounded theory analytic approach.

Literature Review

Many humanities departments and research institutions struggle to integrate digital research methods into their academic programs (Unsworth, et al., 2006). In addition, although archives and libraries are now several decades into the process of building digital collections for humanities scholars, many individual scholars still find it difficult to fully leverage digital materials in their research activities (van Zundert, 2016). Institutional approaches to improving DH scholarship have focused strongly on increased funding for upgrading infrastructure and support services, but this increased investment has not produced a uniformly corresponding boom in DH scholarship (van Zundert, 2012).

One logical approach in institutional efforts to improve DH scholarship is to focus on meeting the *most common* needs of humanities scholars pursuing DH research activities. However, as already discussed, discovering commonalities is a challenge. There is one common denominator, however, that stands out in a literature review on DH scholarship: the need to enable collaboration across highly dissimilar academic disciplines and professional skill sets. DH is an inherently inter-disciplinary activity, requiring a very diverse set of skills. It is rare for a single person to possess the mastery needed in the full spectrum of skills required to complete a DH project. Therefore, if collaboration is a key, then finding a model to study the relationships between the collaborators is critical.

In her introduction to the book, *Cultural Heritage Infrastructure in Digital Humanities,* Agiatis Bernadou argues that,

> The ideal digital Research Infrastructure today should be conceived of primarily as a scholarly ecosystem…(and) to engineer an infrastructure as a sustainable and effective ecosystem calls, therefore, for an understanding of the practices and needs of scholars, archivists, technical specialists as well as other end users of the knowledge production, reproduction and dissemination process (2018).

Joris van Zundert similarly conceives of DH as a "trading zone" similar to an economic market (2016). If we adopt this framework, we can begin to analyze the needs of each group in the ecosystem or market independently from the others without needing an exhaustive understanding of the system as a whole in order to make improvements. Assuming that a market is a defined set of individuals or groups, each exchanging resources (supply) to meet needs (demand) while following some set of rules or norms, then impediments and dysfunctions can be analyzed using the classic tools of costs and incentives (Encyclopaedia Britannica, 2017). Facilitating successful collaboration becomes a matter of defining the groups in the market and understanding each group's resources and needs. Once the boundaries have been defined, it then becomes possible to explore different incentive and support infrastructures to facilitate healthy relationships, efficient communication, and exchange between the various groups. (It is important to note, that by recommending this framework of analysis, this paper does not mean to recommend a neoliberalization of DH, but rather an evaluation of the incentive structures involved, leading to actions on the part of leadership and funding bodies to appropriately adjust those structures toward progress and equality.)

If we expand Bernadou's description into a model, we can identify four important domains in the DH research infrastructure. These are:

1. Data Science Domain
2. Library Services Domain
3. Scholarly Analysis Domain
4. Policies and Funding Domain

Figure 1 shows these domains as concentric circles, with the humanities scholar depicted at the center of the model in blue. In order to accomplish research, the scholar must access original source data depicted as the outer green layer. Between the scholars and the source data is an intervening layer in orange where access is mediated and technical services are offered. All of these domains are supported by the larger domain of policy and funding administration depicted in yellow.

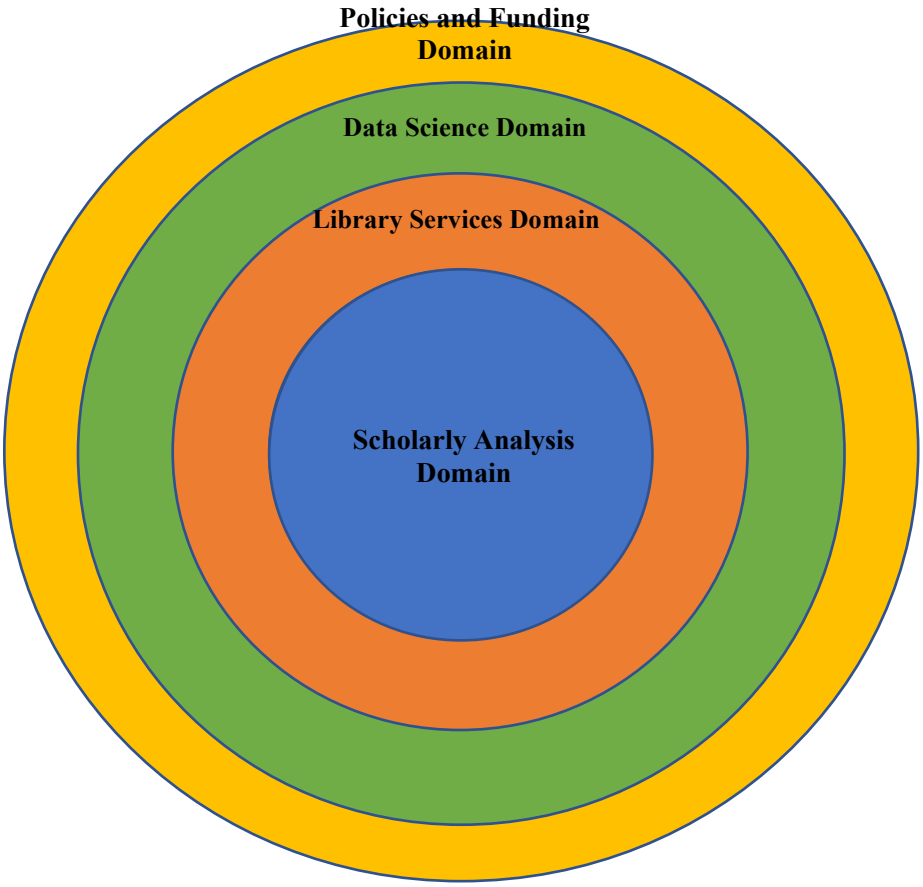


Figure 1. DH Market Domains Model

Each of these domains is populated by people with particular resources, skill sets, and needs which we have labeled as issues in our model. Below, Figure 2 depicts the domains as squares, whose color-corresponding diamonds and circles represent skill sets and issues, respectively. The remainder of the literature review will provide a detailed description of each domain and its associated skill sets and issues, as well as how each domain is connected to the others. We will begin with the data science domain and work our way inwards, leaving the final outer ring, the policies and funding domain, for last.

Figure 2. DH Market Domains with Associated Skills & Issues

1.1 Data Science Domain

The building blocks of any digital humanities project are data. This data is comprised of digital collections from many sources; some born-digital, others digitized from original formats. The digital format opens the door to new approaches to scholarship and analysis not previously possible. Thomas Padilla, provides a useful introduction to understanding collections as data in his article, *On a Collections as Data Imperative.*

> To see collections as data begins with reframing all digital objects as data. Data are defined as ordered information, stored digitally, that are amenable to computation. Wax cylinders, reel to reel tape, vellum manuscripts, websites, masterworks, musical scores, social media, code, and software in digital collections are brought onto the same field of consideration. The value of such a shift can be explored in part by asking how thinking about an object as data multiplies and/or extends the questions that can be asked. For example, if the notion of a single digitized text is shifted from a surrogate of a bound paper object to consider the possibility latent in a form that is computationally processable at the level of thousands or even millions of texts, a move is made toward meaning making that engages affordances unique to data…Meaning making with data is not solely a consideration of whether a computer can be used to process, visualize, and mine them. An orientation to collections as data is about cultivating perception that pushes past the surface of the things that inhabit digital environs (2016).

As Padilla points out, one of the biggest values inherent to using collections as data lies in the exponentially increased scale of what is possible. Instead of individual scholars reading through individual documents one at a time, in the digital realm it's possible to "read" through thousands of documents in a relatively short amount of time. An excellent example of the type of project that is possible in this context is one of the winners of the 2016 National Endowment for the Humanities "Chronicling America" Challenge: *American Lynching: Uncovering a Cultural Narrative,* by Andrew Bales. The project integrated data from the Chronicling America API of over 100 years of historical newspapers available through the Library of Congress, as well as lynching data from

Tuskegee University, and case files from Project HAL to create an interactive chronological map of victim reports and their state-by state distribution (Bales, 2016). This scale of information collation would be extremely costly and time consuming with a traditional, non-digital analysis approach. Using the digital approach produced a methodology and scholarship that was profoundly innovative and provided fresh insights into important historical events.

The massive undertaking of digitizing collections and making them available to scholars as data has primarily fallen to large institutions such as universities, libraries, corporations, and governmental bodies – i.e. the traditional keepers of the cultural and social record. The reasons for this are fairly obvious; the costs and complexity associated with creating and maintaining large archival collections in digital form are no less prohibitive than they are in the traditional formats. In addition, records in any format are of little value without authoritative verification of authenticity and provenance. However, this reality introduces a dilemma. Padilla, summarizes it nicely,

> Power and control resides [sic] in the hands of those that take a data first, representation second mentality – namely corporations, governments, researchers, and increasingly as Bergis Jules has noted – law enforcement agencies. In simple terms, a collections as data imperative entails developing the means to…engage critically with the traces of human activity we collect in the fullest manner possible, native to the complexity of their form, and critically attuned to the possibilities *and* perils that come with their use (2016).

A foundational reading for this paper was the book, *Processing the Past: Contesting Authority in History and the Archives* by Francis Blouin Jr. and William Rosenberg (2011). The authors provide a comprehensive overview of the evolution of the relationship and growing "divide" between the discipline of archival administration and the academic disciplines that make use of archival collections. They make a compelling

argument that the introduction of digital technology necessitates a re-examination of this relationship and how each side understands both their own and the other's role. While scholars have been moving towards a postmodern historiography and shifting away from essentialist Rankean philosophies, the pressure of ever-increasing volumes of records and digital technologies have pushed archivists towards modern records management practices that inherently embrace an institutionally-based essentialism. In short, due to simultaneous changes in historiography and archival administration, these two fields, which were traditionally linked and broadly familiar with one another's scholarly paradigms, are now largely incognizant of each other. The authors even go so far as to posit that, "any visit by a historian to an archival institution is now an exercise in interdisciplinarity (2011)."

Blouin and Rosenberg argue that one of the primary elements of the divide between scholars and archivists is the false perception by many scholars that the archive is a static warehouse of documents and records waiting in something like primal sterility to be discovered, interpreted, and connected to the broader historical narrative by trained historians, while the archivist is simply the neutral caretaker and guardian (2011). In reality, profoundly significant decisions, both additive and subtractive, are made by the archivist before the historian ever arrives, that literally define the historical record. Far from being passive or neutral, archivists are active agents in creating, structuring, and mediating the historical record. The authors quote Jacques Derrida's *Mal d'Archive* stating, "archivization produces as much as it records the event" (1995).

This disconnect is only compounded in the digital context. The layers of decisions, assumptions, and hidden social structures embedded in any given archival record are

multiplied by the exponential power of digital technologies. The breakdown of scholarly communication between researchers and archivists can lead both parties to make incorrect assumptions based on missing data. Blouin and Rosenberg advocate for "bridging the divide" by a deliberate effort on both sides to communicate and educate.

> Digital archives currently allow no point of entry into the principles and practices underlying their formation, other than what the computer screen itself reports…The common notion that historians can understand these practices intuitively is a recipe for poor scholarship. To write good history, scholars must bridge the archival divide by acquiring specific knowledge about archival processes, reading their archives as well as their documents, and by once again understanding what their now distant archivist colleagues actually do. In important ways, this approach constitutes a new historical methodology in which the 'source of the source' is carefully considered and explicated (2011).

It is one thing to say that a new methodology is required and that lines of scholarly communication must be established. It is another to actually put it into practice. In his overview of the process of constructing collections as data, Padilla emphasizes that a key component to success is a focus on "legibility."

> To make collections as data usable, the processes by which they are established must be made legible. These data are the product of design decisions whose purposes are typically not available for a user to consider. Lack of availability can be traced to a predominant understanding of digital collection use that does not address the needs of users who desire to work with collections computationally. The result is the presentation of seamless digital collections that aim to support interactions with objects rather than with the data that comprise those objects. Libraries do not often provide access to the scripts that generate collections derivatives, access to processes for cleaning or subsetting data, access to custom schema that have been used, indications of how representative digital holdings are relative to overall holdings, nor is the quality of data typically indicated. Libraries do not typically expose why some collections have been made available and others have not. Libraries do not typically identify the library staff personally responsible for modifying, describing, and creating collections – a dimension of provenance that must be accessed in order to determine data ability to support research claim. Collections as data possibility is contingent on integrity vouchsafed by expanded documentation practice (2016).

Ultimately, the foundation for successful DH projects is transparent and thorough documentation of the data creation process. Scholars must be able to trace and verify not

only the origins of the data, but the additive and subtractive decisions that define it and shape its format. The non-neutrality of archives is a critical issue for DH scholars. As Hope Olsen points out in her article, *The Power to Name: Representation in Library Catalogs,* a database or archive is an implicit model or hypothesis, with in-built biases and power structures. Archivists and database developers have a responsibility to mediate their work and provide insight into the collections and the processes by which they are digitized and turned into data (2001).

1.2 Library Services Domain

The relationship between humanities scholars and the library services domain is defined by the need for mediation of data access and technical skills. Scholars have limited time and resources to navigate the process of data and digital-skills acquisition. They depend on key experts in libraries, archives, and digital labs to help them navigate the process of finding, accessing, and using complex data sets and analysis tools. William Dudley Pascoe states in *Rapid Bricolage Implementing Digital Humanities,*

> Digital humanities requires expertise that crosses many fields from specific humanities disciplines to software development and production management. DH has a broad range in scale – from a scholar learning basic programming to hack a Python script, to multi-institutional collaborations on neural network learning. Few people are experts in all these fields meaning DH is often a collaboration. The requirements for any individual DH project can differ greatly also requiring IT skill sets that may not be easy to find in any one individual (2018).

Many institutions respond to the need for mediation by trying to construct large, automated infrastructures. A significant body of research exists, particularly in DH circles, examining various approaches to "research infrastructure." Sheila Anderson presents a useful overview of the various schools of thought in her article, *Thinking Infrastructure: What Are Research Infrastructures? (2013).* She points out that many

institutions and funding bodies make the mistake of focusing too much on the infrastructure and too little on the intended users.

> Despite the references to people and organizations, there remains…an idea of infrastructure as a thing with a subtle and underlying discourse of the material nature of infrastructures…Despite the well-made arguments put forward to define systems as inclusive of social, political and economic factors; nevertheless, talk remains about moving from systems to networks to internetworks with the primary focus on technology. The debate is more about a question of control, centralization versus federation, heterogeneity versus homogeneity, from bounded systems to an internetwork, all of which emphasizes the material technological components of the infrastructure…Infrastructures in this context are also expensive, hard to change, and with an air of irreversibility about them. The builders of this infrastructure are the software engineer, the systems analyst, and the developer. For the scholar on the outside looking in, the problem is that much infrastructure development has had the look of 'we are building it and you will come (2013).

While the technology-first, centralized approach has obvious appeal from a funding and user-data tracking point of view, it has proven to be largely unsuccessful at fostering lasting progress in DH. Joris Van Zundert writes in his article, *If You Build It, Will We Come? Large Scale Digital Infrastructures as a Dead End for Digital Humanities (2012):*

> These tall big bulky structures will be waiting for a horde of uniformly behaving humanities scholars that will never come…It is nearly impossible to establish what a generalized infrastructure would look like for high-end innovative projects geared towards humanities research – the sorts that involve experimental pattern detection, large scale analysis of noisy data, and exploratory knowledge visualizations. This near-impossibility follows from the experimental character of the research. The uncertain and volatile nature of innovation determines that it is hard to establish the forms and requirements of any underlying technology or infrastructure. *Innovation by definition is the exploration and investigation of that which is unknown by doing and experimentation. Thus, if the large infrastructural projects are concerned with innovation, the question that they must pose to themselves is: how do we deliver an infrastructure for something that is unknown? And how do we cater to unknown research questions?* (emphasis added)

A big part of the difficulty of catering to the unknown research question is that the users accessing the "infrastructure" are usually experts in their fields and highly specialized. Their specialization and expertise are what empower them to achieve

innovation and creativity, but it is also what make them a particularly difficult set of users to serve. Ulf Thomasen and Peter Kastberg capture the essence of the paradox in their article, *Increasing Specialization of Knowledge Leads to Increasing Fragmentation.*

> To organizations then, there is a direct connection between the degree of specialization and the value potential of connected products or services. The more specialized knowledge becomes, the risk of imitation decreases and the chance of significant marked impact increases…the definitions of specialized knowledge do, however, hint to a significant challenge that is ever-present when working with specialized knowledge: the challenge of fragmentation. As knowledge becomes increasingly specialized, it becomes less common, and therefore less immediately understandable outside of its domain. Specialized knowledge is highly contingent on the domain within which it was created and sanctioned as well as on the particular perspective, expertise, and professional background of its knower(s) (2015).

Assembling the fragmented, specialized expertise of various individuals into a cohesive team to produce a single product or goal, is one of the primary challenges of all organizations in the Information Age. To produce innovation, people must specialize, but specialization isolates them. To capitalize on the advantages gained through specialization, it is necessary to build bridges between specializations that allow people to efficiently re-connect. The question then is: what do those bridges look like?

Researchers Li Lu and Y. Connie Yuan's study, *Shall I Google It or Ask the Competent Villain Down the Hall? The Moderating Role of Information Need in Information Source Selection,* provides insight into the question of how to provide support in highly ambiguous information environments. Specifically, one of the parameters of their study examined the question of how individuals choose between relational (other people) and nonrelational (repositories, search engines, etc.) sources. They state, "When information need increases, our results show that people are more likely to consult relational sources…that people favor oral over written information under

the circumstances of high uncertainty. All these results support earlier findings about the importance of interpersonal communication in information seeking" (2010).

Carol Kuhlthau's well known research on the Information Search Process also provides insight. Her paper, *Information Search Process of Lawyers: a Call for 'Just for Me' Information Services*, provides a useful framework for unpacking the question of how to build research support infrastructure for highly-specialized experts working through a creative and ambiguous process to produce new knowledge (2001). Her research also finds that individuals working through a creative process prefer relational vs. nonrelational sources.

> For the most part, systems have not been successful in identifying information beyond the obvious and conventional. Library and information services, as well, have been more successful in meeting routine information needs than those related to the creative process involved in the more complex work tasks. Most systems and services have been inadequate for supporting the information process that enables workers to create individualized approaches that add value to an enterprise…Task complexity is emerging as an important, influential factor in information-seeking behavior in work situations. More complex tasks are nonroutine, unanalyzable, involve processing equivocal information and evoke different approaches to information seeking than do tasks that are routine, analyzable, and involve less equivocal information.

Her insights into the affective aspect of the information seeking process are key. She finds that experts working towards new and innovative knowledge construction view uncertainty and complexity as a positive sign of progress.

> When information seeking is viewed as a process of construction, an information system and service are needed that go beyond provision for seeking and gathering to support interpretation and use…Research is also showing that uncertainty is viewed not as something merely to be reduced but as indicative of the engagement of the user in a complex problem that requires time for construction (Kuhlthau , 2001).

Her study participants indicated that the nonrelational sources of information were inadequate to meeting the needs of the information seeking process in a complex and creative context. Instead, they turned to their relational sources as the more natural fit.

> All of these lawyers used some type of assistance in information seeking and use to accomplish their work. Assistants were identified as secretaries, paralegals, one used a research assistant, and a few mentioned using a librarian…In many instances, the lawyers hoped or expected that computer systems would be developed that would assist them in their information seeking and use and better accommodate their work. But at this point they expressed some disappointment and reservation about the application of current systems for meeting their information needs, particularly in more complex tasks.
>
> …
>
> Efforts to personalize systems have had limited success with users frequently expressing disappointment in the result. For the most part, these systems have been built from a systems' view of what the technology can do rather than a users' perspective of what the work requires…Most designers have been unable to capture the various ways that users actually accomplish the wide range of tasks that encompasses their work. Therefore, most systems intended to be personalized have been built to support routine tasks and do not respond to the more complex aspects of work.
>
> …
>
> The participants in this exploratory study indicated the need for 'just for me' information systems and services. 'Just for me' incorporates 'just in time' and 'just for you' concepts but goes beyond to provide personal information mediation. 'Just for me' services and systems would be grounded in a clear understanding of an individual's work, the different types of information needed and the range of access to accomplish a variety of tasks (Kuhlthau , 2001).

Scholars working in DH are specialists among specialists. They are not only trying to move forward in knowledge creation within their topical area of expertise, they are also trying to innovate in their methodological approach. They need to be able to access support not only in locating and interpreting information and data, but also in sorting through the technical approaches and how to most efficiently acquire the necessary skills. Even the most basic DH projects demand the combination of advanced ability in multiple areas of skill and expertise, and since extreme polymaths are not a common form of

scholarship, they are not a scalable model for progress. Instead, institutions must focus on building bridges between areas of expertise.

As we have seen, the research suggests that the most effective "bridges" are relational. The research infrastructure most compatible with the process of innovation and knowledge creation is a relational one that gives scholars access to consultative services to help them through the highly complex and ambiguous process of moving the unknown into the known. Institutions interested in building effective research infrastructures to support sustainable progress in DH should focus on building relationships and on the question of "who" rather than the question of "what."

1.3 Scholarly Analysis Domain

Some of the most critical relationships DH scholars must negotiate are the ones within their own discipline. Arguments continue as to whether DH should be considered a methodological subset of traditional humanities fields or an entirely separate field unto itself (Cordell, 2016). Either way, DH practitioners are faced with a number of fundamental challenges in establishing and defining the boundaries, standards, and methods of their scholarship. Understanding the challenges inherent to epistemology in DH is critical to understanding the challenges of collaboration in the broader market.

Most of the issues humanities scholars face in conducting digital research are rooted in the lack of epistemic infrastructure for considering documents and other archival materials to be data, and for thinking of the creation of databases and code as scholarship. In the case of archival materials as data, it is important for humanities scholars to develop a deeper awareness of how these sources are constructed and to engage with them more critically, not just in terms of their provenance and content, but also in terms of the

process by which they were accessioned, described, classified and, if not born digital, digitized. Members of a panel on transparent data research in DH at the 2018 International Digital Humanities Conference stated,

> Preparing data for analysis requires interpretation and is therefore inseparably part of research and should be incorporated into the disciplinary methodology. This calls for an extension of usual source criticism with more specifically digital source criticism. In a typical research project, involving digital data, they are processed with a variety of tools that change them in many ways, making tool results and data at times inseparable. Tool and data criticism are therefore intertwined (Hoekstra, et al, 2018).

In order for scholars to effectively use archival collections as data, they must understand the creation process of that data, and develop a scholarly standard of critical engagement to ensure accurate and rigorous results.

Similarly, scholars must also develop standards for engaging critically with the software tools and computer codes they use to analyze the data they extract from archival collections. Essentially, they must develop new customs to describe and legitimize their methods of analysis and include them as a standard section of all published research. Though not standard practice in the humanities, such methodological analysis is familiar to anyone in the natural or social sciences. In their comprehensive overview of the world of DH, Anne Burdick and her co-authors described the need for new standards of engaging with tools and methods.

> With the rise of new authoring platforms and collaborative environments, 'supporting' apparatuses have been exposed as anything but transparent and neutral, as they not only determine modes of interpretation and navigation, but also condition and guide the production of meaning…Therefore, a whole new set of evaluative questions needs to be asked…These kinds of questions interject a different set of evaluative metrics into humanities scholarship while raising the bar for digital work. We are still at the very earliest stages of understanding and legitimating these emergent knowledge formations. We do not want to lose sight of the core values by which scholarship is judged, and we also want to be sure we can answer skeptics ready to assert that the Digital Humanities is all technique and lacks content (2012).

Ultimately, knowledge theory formation in DH is a paradox. In a 2018 article, James Malazita, professor of Science and Technology studies at Rensselaer Polytechnic Institute, presents a fascinating argument regarding the difficulty of conducting research that mixes computational analysis and critical thought at the same time. He outlines the instrumentalist approach axiomatic to epistemic infrastructure in STEM and how this clashes at a fundamental level with the interpretive approach of the humanities. The difference between these two approaches to research is important to understand because it reveals the central paradox of DH as a marriage between diametrically opposed models of understanding the world. The digital realm is rooted in a formal logic and positivist conceptions of scholarship, while the humanities are grounded in critical thinking and interpretivism (Edel, 1987).

The difference in underlying approach to analysis that exists between STEM and the humanities produces another important disparity between the two fields. Bernadou captured it well in her introduction to *Cultural Heritage Infrastructures in Digital Humanities,*

> A considerable body of work in the humanities is often differentiated from research in the natural sciences by its interest in the particular: a concrete work or corpus, a historical event or period, a culture, an artefact, or an artist, to name some examples. In this light, humanities research can, in many disciplines, be characterized as often being idiographic, aiming to capture an adequate account and provide understanding of a particular phenomenon, rather than nomothetic, aiming to produce generally applicable (and replicable) laws, or law-like generalizations. It is also distinctive in the higher degree of subjectivity, and lower degree of repeatability of research findings. A related consideration, crucial to the construction of knowledge in humanities research, concerns the centrality of recorded information, exemplified in its reliance on the construction and study of homogeneous corpora (of texts, archival resources, visual representations, etc.) and a variety of other, often complex and heterogeneous collections of information objects representing the record of human experience and knowledge (2018).

Humanities scholars' reliance on homogeneous corpora produces particular obstacles when these collections are digitized and employed as data. Alison Langmead, et al. in their study, *Towards Interoperable Network Ontologies for the Digital Humanities,* provide several spectacular examples of the ways in which data in the humanities can be incredibly complex and heterogenous. They describe the difficulties of standardizing early modern dates,

> Early modern temporal data is messy and difficult to record in a commensurable fashion between projects without hegemonically imposing ahistorical standards and losing key information about how data were originally structured. Even if we restrict ourselves to looking at temporal data associated with early modern Christian Europeans – ignoring the Jewish, Islamic, and Chinese calendars, among others – we encounter a variety of often-contradictory systems for encoding dates (2016).

Overcoming this complexity requires DH scholars to engage tools and methods that are unfamiliar to their traditional roots and frequently beyond the capacity of their domain expertise. Cross-disciplinary collaboration of some type is absolutely necessary.

However, collaboration is relatively unfamiliar in traditional humanities. The focus on the particular and on the study of recorded information produces a culture of individualistic scholarship. Roy Rosenzweig points out that an important challenge in DH scholarship is negotiating the cultural gap between the autonomy of traditional humanities and the collaboration of DH. He states, "The singly authored work is the standard for the profession" (2007). He cites statistics that at the time of his study only six percent of the scholarly works indexed in the *Journal of American History* had more than one author, and fewer than two percent had three or more. In contrast, he states, "I would be hard pressed to come up with a digital work of any significance that is the product of a single author" (Rosenzweig, 2007).

To succeed, DH scholars must develop new standards of engaging critically with archival sources, build the epistemology and publishing standards of software and code as scholarly tools, resolve the tensions between old and new methods of humanities scholarship, and introduce new practices and standards of collaborative scholarship across disciplines. Without simultaneous successful adjustments to all of these areas, the long-term development of DH scholarship will continue to stall, and it will be increasingly difficult to attract and retain new DH scholars.

1.4 Policies and Funding Domain

The fourth and final domain in our model of the DH market or ecosystem subsumes the first three under institutional authority. Moreover, it is the source of structure and funding upon which the others depend. The previous sections of this literature review have worked to define each of the groups in the market and to build an understanding of each group's resources and needs. The conclusion in each case has been that success depends on facilitating collaboration, healthy relationships, and efficient communication within and between the various groups in ways that address the particular dysfunctions of each group. While the immediate solutions to dysfunction are often straightforward in terms of identifying desired changes in participants' behavior, the process for connecting the desired behaviors to adequate incentives are not as straightforward. From the leadership perspective, this is what is known in economics as a "principal-agent problem."

> The principal-agent problem is a conflict in priorities between a person or group and the representative authorized to act on their behalf. An agent may act in a way that is contrary to the best interests of the principal…In all of these cases, the principal has little choice in the matter. An agent is necessary to get the job done. However, there are ways to resolve the principal-agent problem. The onus is on the

principal to create incentives for the agent to act as the principal wants (Chappelow, 2019).

DH research isn't possible without collaboration, but successful collaborations are tricky to build, harder to scale and maintain, and impossible to generalize. Knowledge-based institutions as the principals must use scholars, archivists, librarians, technical consultants, and others as their agents to achieve innovation and progress. However, the incentives for each of the various agents are frequently not aligned with the stated goal. Success, therefore, is dependent on careful adjustment and calibration by leadership to the relationships and incentives among the highly unique agents involved in a given institution. Rogers Hollingsworth, in a 1984 article for the Bulletin of American Scientists, *The Snare of Specialization,* captured the problem:

> The establishment of interdisciplinary centers to carry out basic research will not be easy to bring about, given the American university's high degree of institutionalization. First there is the lack of intellectual esteem which academicians have for those in other disciplines. This is frequently exacerbated by the dynamics of interdisciplinary activity when it is tried. Knowledge tends to flow from disciplines which are more theoretically deductive and which have more rigorously organized knowledge and better quantitative methods. In other words, knowledge tends to flow from the harder to the softer disciplines, partly because the harder the discipline, the easier it is to demonstrate the validity and significance of an argument…But more fundamentally, social institutions are very conservative and difficult to change. In the social system of science, the leaders are by definition the ones who have excelled in the present system and tend to be quite incapable of appreciating a different type of system. For this reason, those who are leaders in their respective academic disciplines can hardly be expected to take the lead in promoting interdisciplinary communication. Moreover, since universities are facing a fiscal crisis, it is exceedingly difficult for administrators to shift funds from academic departments to interdisciplinary programs. Similarly, scholars seeking funding for their research will probably choose to play it safe by working in areas in which they have already established their expertise, instead of venturing into new areas.

This perspective makes clear the naivety of hoping DH projects and methods will naturally develop if given enough time. To get different behaviors, one must change the

structure of the rules and rewards. An important panel discussion took place on this topic at the 2018 International Digital Humanities Conference titled, *Precarious Labor in the Digital Humanities.*

> Although stakeholder enthusiasm for digital humanities may be considerable, institutions are still learning how much and what sort of work is necessary to bring a project to completion effectively, and sustainably, and without considerable exploitation. Even supportive and practical administrators may find that they vastly underestimated the work involved – but cannot provide more needed support without more evidence of success in the form of finished projects. Such behavior places an overemphasis on the product rather than the process. Both building a project and developing the skills, policies, and partnerships needed to sustain energy and activity around it take significant time and training (Boyles, et al.).

In other words, if the measuring stick of a successful DH program only counts completed projects and splashy activities or products while ignoring the effort involved in developing robust human networks, meticulous datasets, or shared taxonomies and standards, then it discounts the vast majority of the labor and progress occurring in the background. The *Precarious Labor in the Digital Humanities* panel discussed the issue in explicit terms:

> The labor involved in attending to the needs for community outreach, interface design, user experience, accessibility, and other factors essential to making the metaphorical bridges materialize between these projects and their desired audiences is often precarious, underpaid, or even missing completely from the planning and implementation stages of these projects.
> …
> The central goal of this panel is to initiate conversations about these 'deeply divided forms of human labor' in the digital humanities, often neglected in favor of creating more DH projects. These divisions take many forms: the lack of ethnic, cultural, and economic diversity among DH practitioners; the contingent nature of DH positions; the exploitation of digital laborers within and beyond classrooms; and the challenging or outright dismissal of the value of digital humanities scholarship by tenure and promotion committees.
> …
> Alex Gil has identified the 'miracle worker' as a particular kind of digital laborer, one who is expected to cover a range of roles, responsibilities and projects with a minimal amount of resources, support, and compensation. Miracle workers are

> expected to be competent scholars, accessible tech support, patient project managers, and more…due to a lack of human and fiscal support, the 'miracle worker' often is unable to challenge traditional modes of scholarship that may be ineffective or even harmful…collaborative, public-facing, and iterative digital scholarship proves challenging in environments that encourage 'miracle worker syndrome; because they tend to privilege the monograph at the exclusion of digital work (Boyles, et al., 2018).

Building a "research infrastructure" to support DH is contingent upon being fully aware of all of the different types of labor involved in the process, who is performing that labor, and whether or not they are receiving appropriate acknowledgement and career rewards for their contributions. Because DH is so dependent on a collaborative work model for success, the key to long-term success is the diligent cultivation of healthy relationships between participants and vigilantly rooting out inequalities in labor/reward structures. The environment a DH "ecosystem" needs to thrive isn't necessarily expensively funded or elaborately organized, but it is intensely focused on building bridges between specialties and rewarding all forms of labor in proportion with their effort and value. In their abstract, *Expanding Communities of Practice: The Digital Humanities Research Institute Model,* for a panel presentation during the 2018 Digital Humanities Conference in Mexico City, Lisa Rhody, et. al stated:

> Digital humanities skill development cannot be automated; it is resource intensive. It depends upon a limited number of people to deliver highly personalized training to relatively small cohorts of scholars – a model that is difficult to fund and harder to scale.

From this perspective, no two institutional "research infrastructures" will look exactly the same, since the individuals and relationships involved will be entirely unique to each institution. The process of development will be an iterative one, with all the frustrating awkwardness and messiness inherent in the development of any human relationship. But if leadership is willing to protect and reward the process as much as the product,

innovation is possible. Sheila Anderson states in her article, *What Are Research Infrastructures?*:

> Infrastructure becomes *research infrastructure* as part of a process of change, collaboration, and engagement. In these infrastructures, collection-holding institutions act as creators, curators, and bearers of knowledge about their holdings; technical development seeks not only to capture and represent digital information and content but also the processes by which that knowledge is created and continues to be created as it is analyzed and used; researchers act not just as users but also as 'readers,' of both the collection holding institutions and of the holdings, possessing both archival and artifactual intelligence, and weaving narratives based on interpretive and analytical research methods and processes…Infrastructures thus become dynamic not static, retaining the old, but adding the new. They operate as an ecosystem in which the component parts interact, shift, and change in a constant process of engagement, adjustment, and readjustment (2013).

The key thing for policy makers and funding decision makers to remember in the DH process is that it will be neither linear nor easy. Collaboration isn't a shortcut to success. It's messy and risky. Roy Rosenzweig stated in his article, *Collaboration and the Cyberinfrastructure: Academic Collaboration with Museums and Libraries in the Digital Era,* "Collaboration, to put it simply, is very hard work – work not just on the joint project but also on the process of collaboration. The second biggest mistake after ignoring the possibilities of collaboration is to assume that it will be less work" (2007).

Analysis

1.1 Introduction & Methodology

To further investigate the current needs of DH scholars, interviews were conducted with seven individuals from five universities. These individuals were chosen based on their experience in conducting digital humanities projects. Potential interview subjects were located through the recommendation of their peers or through their published research.

The research was constructed as a qualitative study following a semi-structured interview format. The interviews were conducted over a period of twelve months, each interview was roughly one hour long. All interviews were digitally recorded and then transcribed. The results were then coded using the Nvivo software suite using an inductive grounded theory approach to the data analysis process.

> Researchers can use grounded theory strategies without endorsing mid-century assumptions of an objective external reality, a passive, neutral observer, or a detached, narrow empiricism. If, instead, we start with the assumption that social reality is multiple processual, and constructed, then we must take the researcher's position, privileges, perspective, and interactions into account as an inherent part of the research reality…With grounded theory methods, you shape and reshape your data collection and, therefore, refine your data and increase your knowledge…When combined with insight and industry, grounded theory methods offer sharp tools for generating, mining, and making sense of data (Charmaz, 2014).

An original set of eight coded categories was restructured in several subsequent rounds of coding into three overarching categories, each with two sub-categories. The final set of categories is laid out in Table 1 on the next page.

Table 1.

Primary Category	Sub-Categories
Data Science Domain	• **Archivists**: Understanding Collection Creation Decisions • **Database Programmers**: Understanding Database Formatting
Library Services Domain	• **Librarians**: Mediated Access to Data • **IT Techs**: Mediated Tools, Skills and Methods
Administration & Policy Domain	• Boundaries and Responsibilities Among DH Collaborators • Incentives Among DH Collaborators

In the interviews, subjects were asked to consider the following:

Libraries are rapidly converting their archives of humanities-related original research materials to all-digital formats, significantly altering the collection management practices of their profession. Humanities scholars are also increasingly publishing their research output in non-traditional, "born digital" venues (e.g. interactive websites, You-Tube videos, etc.). As these trends occur, how should libraries also modernize their customer service functions (e.g. by standing up "Digital Labs") to provide advanced computational tools and specialized services that help humanities scholars exploit the digitization of humanities scholarship?

Subjects were particularly asked to consider some of the following questions:

a) Describe the overall process you have followed in your humanities research to access and analyze original source materials.

b) Describe any especially helpful customer service functions provided by archivists or librarians in the special collections you've worked with in your research

c) Describe your impression of the ongoing trend of Libraries converting their archives of original research materials to all-digital formats

   i. How well is the conversion to digital collections being carried out?

   ii. What impact is this trend currently having on your research activities?

   iii. What impact do you think it will have in 5 years? 10 years? On the next generation of scholars?

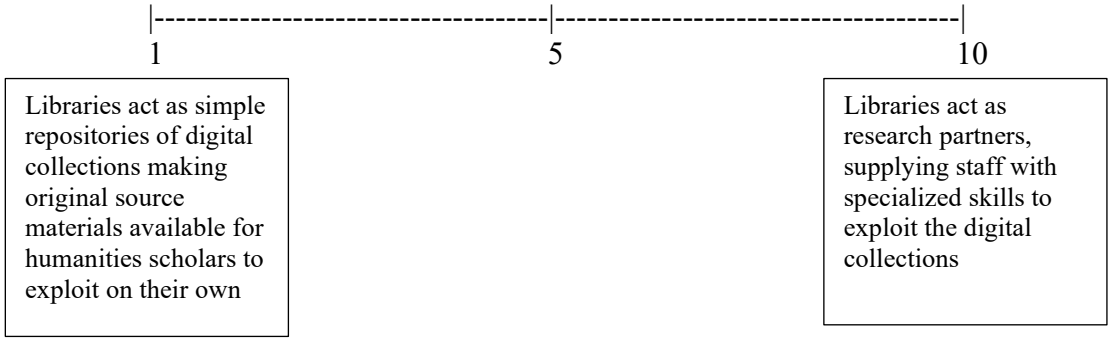
   iv. What concerns do you have?

   v. How might the conversion be a positive or negative development for the future?

   vi. How is your data analysis changing with the advent of digital special collections?

   vii. Is your research at the stage of needing "big data" analytic skills?

d) As the trend toward the digitization of humanities scholarship continues, describe how libraries could best modernize their customer service functions. In your description, mentally consider the scale below. From 1 to 10, where should libraries sit in their customer service model for digital special collections?



e) What specialized digital analysis software TOOLS are most important?
    i. Who should supply these specialized digital tools?
    ii. To what degree should scholars and/or their individual departments be expected to provide their own specialized digital tools?
    iii. To what degree should the libraries be expected to provide them? Some combination? What would work best, and why?
f) What specialized computer SKILLS are most important?
    i. Who should supply these specialized skills?
    ii. To what degree should scholars and/or their individual departments be expected to provide training on these new digital research skills?
    iii. To what degree should the libraries be expected to provide them as a customer service? Some combination? What would work best, and why?
g) If librarians provide extensive digital and analytical support, should conventions be established to formally acknowledge their scholarly contributions? Why or why not?
h) Should libraries function as the ultimate destination for "born digital" scholarly work or should individual departments and scholars remain responsible?
    i. If libraries are the end point, how should decisions be made about which projects get preserved?
    ii. How can channels of communication best be established to allow scholars and libraries to collaborate effectively?

The scholars were also introduced to the model in Figure 3 and asked to provide feedback on its accuracy as a depiction of their perspective of library/archival services. The model is titled “The Fourth Librarian” which is an allusion to the concept of the fourth wall in theater. The graphic roughly describes an academic library. On the bottom floor, we see on the far left general users - members of the public and/or undergraduate students. On the far right, on the bottom floor, is the general collection, comprised principally of secondary sources. The general collection is maintained by cataloging and collections librarians, depicted in the middle column. Access to the general collection is mediated by reference librarians.

On the second floor, on the far left, are academic users – expert users who are on the cutting edge of their research field and creating new ideas, new knowledge. The collection on this floor is archival - composed of primary sources often unique to the institution. Those archival collections, whether in analog or digital form, are created and maintained by archivists, preservationists, and in the digital form database developers. Both lower floors come under the authority of the upper floor – composed of administration and those who control the budget.

The fourth librarian is the quadrant on the second floor with the question mark in it. Traditionally, access to the archives was mediated for scholars by those who created the archives. Scholars and archivists interacted directly. The question posed to interviewees was: With the added complexity of the digital format, do you see a need for a fourth librarian to act as a mediator between scholars and the people constructing the archives? Is there enough of a gap between the knowledge and the skills that humanities scholars

typically walk into an institution with and what is needed to functionally use the archives, to require mediation?

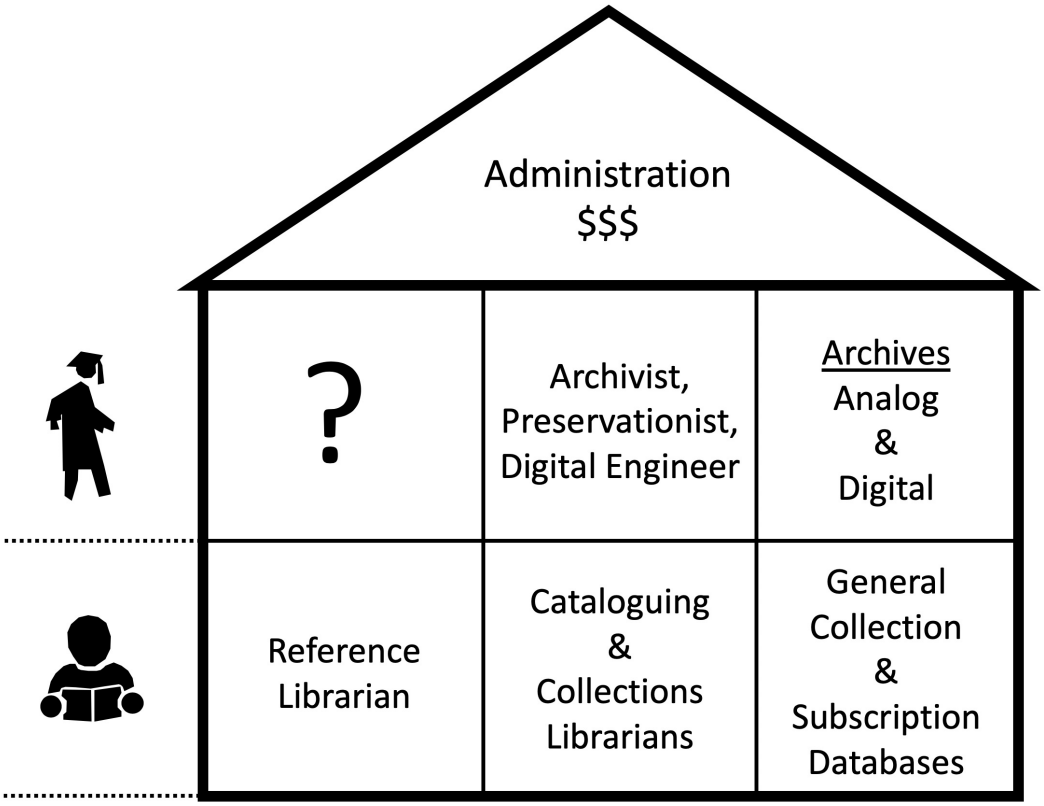


Figure 3. The Fourth Librarian – Locating DH Support in the Academic Library

1.2 Data Science Domain

One of the three primary categories to emerge from the interviews was centered around the issue of the data DH scholars use as source materials. As laid out in the literature review section, much of the data are digitized or born-digital documents collections held by archives or other special collections in libraries. The literature review provided an overview of current discussions among scholars regarding the non-neutrality of archival collections, the hidden labor and influence of archivists in creating the

collections in both analog and digital formats, and the urgent need for humanities scholars to develop a more critical approach to engagement with archival collections and their creators. The interviews conducted for this research paper added two key insights: First, metadata is data. Second, curation of such data is a scholarly pursuit.

Collections Creation

Archivists have been creating collections for all of history; however, the scale of data has changed dramatically in the information age. Interviewee SR states, “I think one of the consequences of digitization is that the humanities researchers are often working with far greater scales of materials than they have in the past, which is making it more pressing for them to be able to organize that material.”

In the past, the material that was included in collections was somewhat arbitrary and was simply based on which artifacts were physically preserved throughout history. DB shares, “It seems enormously accidental sometimes, what resources are available to scholars. There’s an enormous amount of stuff that’s lost! Like the fire of the library in Alexandria, and it’s accidental then, what we have. So, the farther back in time you look the more you are constrained by this history of accidents as to what is still around.”

However, modern technology has brought on an explosive proliferation of data. This new wave of data is wonderful for researchers, as DB states “I need more data rather than less data.” However, DB qualifies this enthusiasm for “more data” with recognition of a new responsibility, “But now, in the information age, we’re having to make a whole lot more decisions about information, what to keep and not to keep. So, there’s a lot less

accident about it." JZ provides a similar warning, stating it this way, "There is too little knowledge about how important the selection of data is, how much data cleaning, as they call it, has been going on when data is finally a digital resource."

The interview subjects repeatedly discussed that someone has to go through a given set of documents and sort them, preserve them, catalog them, label them, and somehow process a pile of undefined documents with the intent to turn it into a well-defined and searchable collection. Inherent in that process is the addition and subtraction of information, adding structure, removing chaos, making assumptions, and becoming intimately familiar with the material in the collection. All of this creates a foundation upon which humanities scholars base their research and scholarship. In keeping with what was found during background reading for the literature review, interviewees agreed that this work has been somewhat hidden from humanities scholars in the past, and many scholars have approached collections without challenging the assumption that they are working with raw and unbiased source material, when in fact it is highly processed and perhaps even biased or censored somehow.

> I'm trying to teach my students in my rhetoric classes how, when they're given statistics that - what questions are asked, what data is included, what data is excluded, are interpretive acts. So, what they think are objective truths are completely hidden through all these layers of interpretation. [GG]

ML reinforces this thought, pointing out how crucial the decision making and metadata surrounding a piece of data is, and how a single piece of new data can completely change all former research and analysis.

> It *does* matter how librarians have set things up as far as metadata. It took a lot of time and thought and effort to do so and it has direct effects on scholarship. One of

> my favorite women poets from the Renaissance, her Biblical epic was attributed to her brother until probably, I don't know, 15, 20 years ago. So, to me that's a big deal, because otherwise, the people who worked to figure out that it was her and not her brother and have been working to change things like the metadata and stuff so that it's searchable, allows me to know that attribution, but it shapes scholarship before then. [ML]

The seemingly small choices of what to name something, what to associate it with, how to categorize it, when the archivist is pulling a piece of data out of the chaos of the maelstrom of information, has a profound influence on its downstream trajectory. It can completely change how it's used, how it's perceived, what it's connected to, who it's aimed at, which audience will have access to it, and it can have both positive and profoundly negative effects. DB ends with this thought, "Folks sit around trying to come up with classification schemes to talk about domains of inquiry. And what they characterize, what they choose, predicates what we're going to use. It defines what mechanisms people think that data characterizes, and who's going to look at it. And that is an enormous [responsibility]."

All of this culminates in a rather stunning statement from SR:

> Digital humanities is making it more and more relevant, how they describe materials, not just how they organize them…*this thing called metadata is actually data*…[SR]

That bears repeating: **"this thing called metadata is actually data**."

Metadata as Scholarship

Within digital humanities, **database programmers** are becoming the modern equivalent of traditional archivists. In fact, many direct parallels can be drawn between traditional humanities studies and digital humanities studies. Which leads to an

interesting question: what is the difference between the two? Is a digital humanities scholar simply a humanities scholar with a computer? The answer lies in the innovation and scale that the digital domain unlocks.

> Right now one of the problems I'm seeing is that a lot of analysis that is going on in digital humanities, but also humanities farther afield, is comparing pears to apples. Because one academic is using an awful lot of data and using computational analysis while the other researcher is reading one source. Then, they're trying to come into the same conversation, which is just not a good match, not a good fit. [JZ]

In other words, using an advanced search tool to find a single document shouldn't necessarily qualify as digital humanities scholarship. This action is really just an augmented version of using the old-fashioned card catalog. The search has simply been translated into another medium without innovating any new methodologies. It's the equivalent of sending a letter by email instead of by post. In the end, the action was still sending a letter. "It's absolutely what is, in many corners of the academy, still going on. It's old changed methods that just remediated into exactly the same form in a digital environment." [JZ]

On the flip side of the coin, doing a lot of advanced programming work, such as creating complex graphical user interfaces or websites is not necessarily digital humanities work either.

> The user interface and flashiness of putting in a particular edition of a website is nice, and I think there's a place for it in terms of engaging the public and engaging the wider audience, but I think at the end of the day, what you think about is, is this dataset any good? … Whatever styling, JavaScript styling, or CSS styling has been done… that's just going to break over time. Those things aren't, to me, the scholarly work. It's what you've done with the underlying [data]. [GG]

So what actually provides scholarly intellectual merit in digital humanities? “One of the ways that I look at it is, I think the contribution that cannot be accomplished by a machine, that a computer itself cannot do, I think that that's part of what provides the intellectual merit.” [DB]

Naturally, with the advent of ever increasingly powerful computational tools, such as optical character recognition and machine learning, there has been a strong desire to integrate these powerful tools into the digital humanities research domain. Some may worry that computers are becoming so powerful that they will soon be replacing the digital humanities scholars completely. However, JZ curbs these worries. “It’s a very fundamental problem. Computers are simply, from the bottom up, not very well designed to do the type of reasoning that scholars are used to.” [JZ]

He goes on to explain that even the most powerful algorithms can’t compare to the human’s ability to reason within the abstract. Especially, when it comes to associating imprecise data, such as is the case with creating a digital collection.

> We have been working, it seems, for ages on statistical approaches and artificial intelligence methods that are reasonable tools to inject into humanities research. But there’s a more fundamental problem, I think, which is that the computer logic that we use nowadays, including your mobile phone, your dishwasher, everything, is still using first order logic. First order logic is basically zeros and ones. It’s Boolean logic. And that’s simply not how humans create and process knowledge. We are very good at associating things that a computer would never associate together. We are very good at using imprecise knowledge; taking things that we are not quite certain of, and still being able to make something useful... [to create] useful inferences from that imperfect information. [JZ]

In other words, computers are not anywhere close to replacing human scholars. Therefore, the computers must be carefully guided by individuals that are experts in their

respective humanities fields in order to produce anything resembling scholarly results. This guidance is provided by careful curation and tagging of the datasets. Which leads us to another stunning quote, this time from GG:

> "The curation of the [data], and marking it up, is the scholarly pursuit." [GG]

Metadata is data, and the curation of such data is a scholarly pursuit.

1.3 Library Services Domain

Library services is a combination of collections creation, data curation, and 'just for me' solutions. These 'just for me' solutions can take a wide variety of formats across all stages of a scholarly project, from the initial brainstorming session to the final draft, frequently focusing on collaborative creative problem solving and finding ways to make information more conveniently accessible for individuals. Impressively, this service model has been successfully functioning for millennia regardless of who might walk through the doors, whether a casually interested individual just beginning a scholastic journey or a deeply knowledgeable scholar with years of advanced research already underway. With the advent of digital data, the threshold of knowledge required for accessing information has steadily been growing higher, thus making the 'just for me' services of libraries all the more critical.

> As the head of Graduate Programs [in the Library Scholars Lab], my job is to try and encourage people to come in who might feel like this kind of work is inaccessible to them. So, I usually tell people that if they just have an interest, we can take them to the next step. They don't necessarily need to know anything when they come in our door. … We are trying to meet people halfway. And we try and accommodate people who already have an advanced research agenda, but other people who are maybe just getting started. And that could be everything from helping them brainstorm a project

> and scope the project, to finding the materials for the project, to collaborating on some prototyping of some software so that they can take things to the next level themselves; or meeting week by week with them and programming side-by-side and teaching them the process by which we are going about making something with the whole time. [BW]

This type of service is exactly what digital humanities scholars are looking for in droves. Whether it's copyright issues, or paywalls, or simply the scale of data available, the challenges associated with accessing said data, DH scholars are champing at the bit to find solutions. "This to me is the hardest part of doing humanities research… Is getting to text." [GG] He goes on to explain how he tried to access copyrighted material for his project, "I was trying to get copies of the Google digitized versions, but that would require UNC to sign an agreement with the Hathi trust and Google… So, we were working with UNC's legal team to get them to sign off on it, and eventually it got to a point that UNC didn't feel comfortable signing an agreement with Google so we weren't able to get those versions of the text." [GG]

Paywalls can also be a blocker. GG explains how for another research project, he was trying to access tweets for a particular hashtag. "You can go to Twitter and you can ask for all the tweets to a particular hashtag, but I think the minimum they charge you is $10,000 for something like that."

Even after a scholar gains access to a particular dataset, the data itself can be inaccessible due to its complexity.

> "There is as much text data out there in the world as you could possibly want, but the problem is that a lot of it is inaccessible in particular ways: either it's, you know, buried in images of file... of manuscripts, or images in text but doesn't actually have… You don't have access to the readable text underneath for a computer to

> process. Or, it's not structured in a way that's… The computer can necessarily get at and parse." [BW]

> "All of the big, collaborative, digital projects have grant funding. That's what allows them to be big, collaborative projects. And yet, I strongly feel we give people the wrong impression of what digital scholarship is, if that is what we hold up as digital scholarship." [SR]

This is why the 'just for me' solutions provided by libraries and digital labs are so critical and valuable to scholars. These librarians are providing creative solutions that allow scholars to break through or at least go around these accessibility walls.

> "There's a librarian at Duke that I partner with, and I'm doing a project on the HB2 tweets that happened last year. And, he was able to use the software they have at Duke for social feed manager to get more of the tweets from HB2. … So, he provided much more information than I could ever find on my own, and that information is pretty much this whole project." [GG]

Unfortunately, these collaborative 'just for me' services provided by librarians haven't been traditionally viewed or recognized as being scholarly. In the past, these services have been considered to be more of a background support role than a principle part of the scholarly work.

> "Simply amassing a body of information, or making that body of information more tractable to computers or to the public, or whatever it may be, is often times thought of as being with the plumbing work rather than the scholarly work." [DB]

However, as these digital humanities scholars are struggling with these challenges and pioneering new methods for scholarly research, they are also pioneering new mindsets about what exactly counts as scholarship. DH scholars are beginning to realize that the individuals who create and curate collections have intimate knowledge of not only how to access the material, but what is inside, how it's organized, and a wealth of valuable insight that directly translate to good scholarship.

> I've tried to write pieces about the struggles I've had with getting to text and working with it. And I had this epiphany about the moment we're in… We're in the worst moment. I mean, we're in a very exciting moment, but were in the worst moment for this stuff. Because, we [DH scholars] have to do double work.
>
> I have to do the work of a traditional humanities scholar plus all the other digital stuff in order for it to be validated. And so, I think the same case is probably going to need to [apply] for the libraries, because, I think scholars just don't know what questions to ask yet. And I think the more services that can be provided, that's only going to, in turn, help scholars find what questions that they can ask of digital material. [GG]

As the work of DH scholars and library services begins to become more and more collaborative and the work of DH scholars begins to parallel that of librarians more and more, the DH humanities scholars are starting to forge a collective recognition of the scholarship of library services. “I think the boundaries of what counts as scholarship are shifting,” says DB. BW states it this way. “The scholars are become more aware of the kind of deep intellectual decisions and labor and scholarly thinking that go in to creating data sets. And to become more aware of their own ability to create them, and then, the librarians get more credit for the kinds of work that they're doing.”

Even if this attitude is not recognized at an institutional level, it’s certainly manifested at an individual and personal level.

> I think there's a point where the librarian provides enough material to be a second author on a paper…He and I have talked… He'll be a second author on that paper, because the method of being able to program something in Twitter to scrape all those tweets… That is such a huge part of the project that I just don't feel like it's right not to include him in the paper. [GG]

In other words, these ‘just for me’ solutions are being translated directly into scholarly recognition. “I like the idea that the fourth librarian is this person who’s explicitly trying to reach across the boundaries between the library and staff in the collections

management side of things and reach out to people who might not necessarily know how to work with the materials, but can meet them halfway." [BW]

Stephen Roberts summarizes it:

> Digital scholars are often aware of, because they're doing parallel work to what librarians do, and wanting to be credited as doing scholarship when they do it. ***That goes with recognizing that that's what librarians are doing as well.*** [SR]

Technical

After DH scholars have gained access to data, in order to conduct research, they then need to process said data and manipulate it in various ways. This data processing usually requires technical knowledge beyond what most DH scholars naturally possess. This means that any facility wishing to support DH scholars must provide a measure of technical support: **Technical Consultants**

> I think that we can't all do everything, but I like the idea of digital labs, especially housed in libraries, if there could be a space that is interdisciplinary. Like I've said, I've really enjoyed collaborating with my friend who's a computer scientist previously. At the same time, I know that I'm not going to have time to learn multiple coding languages probably ever. But, if there is a way of creating a space where we could open dialogs with people who might be undergraduate students in the computer science department or graduate students who maybe are really interested in [humanities]. [ML]

As ML points out, these computer scientists don't necessarily need to be hired, they could be incorporated by creating collaborations with computer science departments or similar. Interestingly, she also mentions that she is probably not ever going to find the time to learn how to write "multiple coding languages" herself, indicating she knows some limited coding, but not as much as she would like. This comment touches on a much broader question: specifically, what level of digital literacy should be expected of

DH scholars? SR emphasizes the importance of this question as well: “I think the real challenge of doing this in the library is the subject expertise is not at the same level at every field.”

Thankfully, DB provides a supremely simple model of the various levels of expertise that can enter a digital lab:

> So, there are three different tiers of technological wherewithal. One is almost none. [An individual] can look at [technology] like they look at it on a piece of paper and that's what the high-end current generation scholars are doing. Then there's graduate students who have some facility with tools and the native capabilities of those tools in order to make it more like what is wanted. And then there's do we need to change the tooling? It’s a different level. [DB]

This model is further developed and presented below in a table below:

Technical Wherewithal	**Description**
Almost None	A high-end scholar, an expert in a given humanities field, with no understanding of what technology exists or what it can accomplish.
Some Facility	Has a basic understanding of existing tools and what those tools can accomplish. Comfortable doing simple things with the tools independently.
Expert	Capable of building their own software tools or using existing tools at a very advanced level.

Although a simple model, this is (perhaps surprisingly) a very helpful basis for understanding further discussion around the topic. For example, what level are most humanities scholars currently? What level should they be at in order to realistically operate a sustainable digital lab? Can we educate a few of them to be experts?

DB hits the nail on the head with this set of questions. “How do you build enough literacy among scholars so that they can have effective conversations with the tech

professionals? I think that goes back to the question about, how much of it should be service work versus collaborative work?"

The level of expertise with which individuals come into a digital lab directly dictates the type of service that a digital lab must provide. Let's looks at the technical wherewithal table again from the standpoint of the service required of a digital lab.

Technical Wherewithal	Service Model
Almost None	Digital lab does all the work
Some Facility	Collaborative environment
Expert	Little to no help needed

When we look at the model from this standpoint, it becomes immediately obvious that the sweet spot for collaborative work occurs when humanities scholars have a basic competency in technology but are not experts. BW, head of a scholars' program, states that is exactly where he aims to be with is lab. "We're not the sort of place where you would drop off a project and we will do it for you, but rather, we'll actually collaborate and engage with you, to the point of…The phrase that the R&D team always uses: you own your own project."

Scholars should have a certain level of digital skills in order to be able to produce this research. However, there is a limit to what can be accomplished by one individual. It's the return on investment for any given scholar in the skills acquisition field. GG states, "I don't think the answer is to teach humanities scholars how to code. …We need to have staff in departments that know how to code and program."

In other words, a lab can't build a sustainable innovation model that depends on everyone being a polymath - an expert in both their chosen field of humanities study as

well as an expert in computer science. There are only so many really high-level polymaths in the world. If a lab is depending on those types of people appearing on a regular basis to move forward in a certain field of research, that is clearly not a sustainable model for progress. A realistic model must allow people to specialize and collaborate in a complex environment in order to be able to continue to make incremental progress.

There are those who would like everyone to be an expert, in fact JZ admits that this is a common flaw in many analyses of DH infrastructures, and goes on to emphasize that it's not even realistic to attempt to train people to this level of expertise.

> I *would like to say* that argument is completely untrue and false. However, I am afraid it's completely right. I know that one fundamental flaw in my reasoning in most of my work is that it suggests that we need these people that are both highly expert scholars and highly expert programmers. In the 20 odd years that I've been active in what you would have called digital humanities in hindsight, I have met maybe 10 people that are like that.
>
> It's an impossible way of trying to come up with a sustainable model for innovation to try to even create people like that because apparently people don't become like that. You can't teach people to be like that. They're really exceptions, which is very unfortunate of course. [JZ]

Not only that, but JZ brings out an additional concern: the very act of attempting to train someone to that level of hyper expertise could actually be detrimental to their career.

> I think a very interesting question that is surfacing is whether the very expert knowledge that such a person would develop doing the project would not also hurt her or his abilities to move to other projects again. [JZ]
>
> I agree that increasingly we need this hyper expertise, which cannot be embodied by one body, by one researcher, one librarian or one expert. However, the expertise itself is so expensive, as it were. If you would somehow instill that, embody that in a specific person, would you still be able to move that person to other projects? In other

> words, very, very high level of expertise that is needed might lessen the ability of these persons to be shopped around in different research configurations. [JZ]

This further solidifies the notion that the necessary skill sets can ONLY be found in a collaborative environment. However, in a collaborative environment, the time and resources of various individuals is limited.

Let's look at the technical wherewithal table once again, from the perspective of population:

Technical Wherewithal	Population
Almost None	The majority of scholars fall into this category
Some Facility	A motivated minority
Expert	Vanishingly small

A digital lab cannot support anyone and everyone. There must be some sort of threshold of entry to use the resources. So, it makes sense to limit the time and resources of the lab to a motivated minority of individuals who have demonstrated sufficient desire to learn a basic competency. On the flip side of the coin, devoting time and effort to experts, who are such a small portion of the population, and arguable don't need much help doesn't make sense either. Thus, it makes the most sense to focus time, tools, and resources on the middle tier of expertise.

GG explains it this way: "There doesn't necessarily [need to be] skill, but there needs to be fluency." People coming into the lab should at least be familiar with the terminology and the art of the possible. He goes on to explain, "I think half the battle for humanists is either they don't think [technology] can do anything or they think it can do everything…We have a lot of scholars that know what they want to do, and they think everything is possible; however, it is not possible. And there's a lot of either, working

really hard to find a solution (and no one realizing how hard that solution is) or just managing expectations…'we can't do that, but what about this?"

Some may argue that providing the right tools can elevate the level at which a DH scholar operates, bypassing the need for a technical consultant or a "4th librarian." However, JZ argues against this mentality, explaining that there is often a "gap", or a mismatch between the provided tools and the users.

> I definitely agree, it's not in providing or buying a tool, software, hardware, et cetera. It's more, there's a gap between the knowledge on the academic side of the story and on the side of the libraries or repositories. This gap of knowledge comes in two ways, weirdly enough.
>
> One is the gap between academics that have mastered the digital skills, which might run from knowing very well how to use digital images of original documents and resources to academics that are able to program tools themselves and are in need of large quantities of digital data coming out of certain repositories. So, you have experts on the digital side offering access to resources that those academics need with access direct to the data level of resources. Usually that kind of academics are confronted with all kinds of nice and very costly graphical interfaces that they really can't work with. They really have to get through those graphical interfaces to come to the real data that they really need.
>
> The other gap, that also is very much in existence is, that you have an awful amount of academics that are really expert in their field, but they don't have any digital skills to talk of really. They are prepared, it's not resistance, it's not wanting to use tools. It's that their skills haven't been trained to use those tools. Usually the argument is that for such kind of academics, the graphical interface is a nice tool because it offers the access to digital resources in ways that is attractive for people with less digital skills. However, I agree that the big problem there is that if you give those researchers just the graphical interfaces, the search tools, the easy search tools, the indexes, that's all they will learn. That's all they use, all the functionality they will be able to get from such a resource. And it's exactly in the role of mediator or fourth librarian, as you call it, to try to work with both skilled and unskilled academics. Especially, I think, to try to offer almost teaching, I would say, to un-digitally unskilled academics to level the playing field for the computational analysis of digital resources. [JZ]

Essentially, he is drawing the table of technical wherewithal from the perspective of tool usage:

Technical Wherewithal	Tool Usage
Almost None	Need simple software with fancy graphical interfaces, but that's all they will learn and they don't actually develop any new research methodologies
Some Facility	n/a
Expert	Fancy graphical interfaces get in the way and are not used. Tend to build their own software tools.

JZ doesn't address the middle tier, but once again, it seems clear that most tools that a digital lab might consider providing should be aimed at the DH scholars in the middle tier.

In terms of what libraries and digital labs should aim a providing: the model that emerges from all of this seems to be that of a Technical Consultant. The DH scholars that come into the lab should have a very basic understanding of programming. This works well both ways as the technical consultants learn a little about the various humanities fields that they work with and the DH scholars learn a little about programming. BW states, "If you actually engage with the people on both ends, and collaborate with them, you get a much better sense of the guts of each project. So, the scholars' lab's general motto is that we want you to get your hands dirty, so that you at least know what dirt looks like.

1.4 Administration & Policy Domain

Boundaries & Responsibilities

So far, we have analyzed the collections themselves along with the archivists who create them. Then, we turned our attention to the libraries and institutions that house the collections and the services they provide. Finally, we looked at training and how to help scholars get their hands dirty. This leaves one final section to address: **administration and policies**.

> I would argue that the level where this knowledge needs to land and really needs to be acknowledged is actually at the administrative level: funding bodies, deans of universities, of faculty programs ... The people that really make the decisions on who is going to collaborate together on this specific project, or on this specific problem; and how are we going to fund them? [JZ]

Identifying the right administrative levers to pull or exactly which policies to change and how to change them is completely outside of the scope of this paper and entirely depends on the goals and ethos of each institution.

> It is a very, very complex problem. It's not a matter of, "Oh, if we simply change this political approach, everything would fall in place." It's much more like so many expert people's policies and strategies are involved that it needs a very, very almost holistic, I would say, approach to start solving this problem. [JZ]

So, in this section we will provide a "holistic" view of what a successful DH lab looks like from an administrative perspective, the challenges they face, as well as recommendations and considerations for helping a DH lab to thrive. Specifically, we will examine structure, funding, project scale, tools, and both the challenges and value of

maintaining digital collections. The bulk of the perspectives on DH lab administration are drawn from a single interview with a subject who was the director of a major DH center.

1. STRUCTURE

To begin with, let's look at what a fully matured digital lab looks like. This will provide a very helpful framework from which to conduct the rest of the discussion. SR, who heads one of the oldest digital humanities centers in the United States, describes the structure of their lab, which has been around for over 20 years and has an accomplished history of award-winning DH projects.

> We're divided into three divisions. The ***education division*** is where the center started. They work with K to 16 teachers and classrooms, and build educational resources and tours as extensions of teaching. That's very much how the center got started. Though, it's increasingly the area of digital humanities that is not as well funded as it has been in the past. We do a lot more work with teachers than most digital humanities centers do, and always have, for 20 years or so. That includes a lot of professional development work, and online professional development courses for teachers.
>
> The ed division also built the online courses for the history departments Graduate Certificate in Public Digital Humanities, which is done in conjunction with the Smithsonian. We have a ***public projects division*** which emerged out of education, with a particular focus on the cultural heritage sector. So, a different kind of audience, a different attention. We do software development in that division, so that's the home of Omeka. Otherwise, that division works with cultural heritage institutions to develop online collections, and does professional development targeted at those kinds of communities and audiences. Then we have a ***research division***, which is the home of Zotero, and is focused primarily on university academics and their needs. It's also the home of Tropy and the home of PressForward. [SR]

The divisions are organized into a table on the next page.

Division	Description
Education	Works with classrooms to create educational resources as extensions of teaching
Public Projects	Online collections with a focus on the cultural heritage sector Software Development: Omeka, a tool for hosting digital collections online
Research	Collaborates with university academics Software Development: • Zotero, a research assistant for managing bibliographic data and related research materials • Tropy, a tool for organizing and tagging photos • PressForward, an online scholarly publications collection tool

Perhaps surprisingly, the lab is not a stand-alone entity, but is actually a division of the history department at GMU. “We're formally part of the history department, which I think people lose track of. They think we're somehow free floating in Mason. We're not. We get a lot of our institutional support filtered through the history department, rather than do it from other parts of George Mason. That includes the leadership of the center.” [SR]

Understandably, the department under which the digital lab is housed directly influences the type of work they do. “The expertise of the scholars that are involved in the center, is very much in history, and not in any other kinds of discipline. So, we do a lot of what we do coming out of our work as historians.” [SR] SR then offers a glimpse into the scale of operations at the lab.

> Across all those divisions, there's about 45 people who work at the center all up. That includes about a dozen graduate research assistants. We're not a fellowship based entity for the graduate students. They work here as research assistants, which means they work on a whole range of our other projects, not on their own projects, as

> much as I know they would like to at various times. That does give them an enormous exposure to both a grant funded institution, and a whole range of different projects. … Not all of those 45 people are based here at the center. We probably have close to a dozen remote contract developers. [SR]

Due to the scale at which they operate, they are very reliant on grant funding which creates constant pressure to find new projects. “Because we're grant funded, we're constantly having to find new projects. … That makes a less stable, high stress, high organization to maintain, notwithstanding the fact that we've been around for 22 years. We do ... we have managed to do that but the center fluctuates a little bit in size, depending on the style of the projects we bring in.” [SR]

2. FUNDING

Clearly, funding facilitates a great deal, but SR provides some warning regarding when and how to seek funding. “The trend has been increasingly to put digital humanities centers, and libraries, and other centers where there's hard money funding. The tradeoff for this, obviously, is there's a lot more service to the institution in that arrangement.” [SR]

He goes on to explain that the leadership of his lab is officially “0.5 faculty.” In other words, they are supposed to divide their time and efforts 50/50 between the lab and the history department, an arrangement that SR says is “completely impractical, because it takes more than 50% of our time to do our obligations here at the center.” [SR] SR goes on to point out that despite the pressures that come with grant funding, it actually provides a certain freedom to those working in the lab (besides the 0.5 faculty).

> For better or for worse, because we get so little money from Mason, we have no service obligations to Mason. We're somewhat bizarrely outward facing from the institution. … We have, because of our grant funding, an increasingly unusual outward orientation and lack of service obligations. [SR]

Not surprisingly, funding is more easily found for the hard sciences. “I don't think we're ever in any danger of getting the kind of institutional support that's provided to science labs, which is also bound up with the amount of grant money they bring in.” [SR] However, because his lab is mature enough to develop customized software, there are increased opportunities for funding. “I think the software development end is increasingly more something that makes us distinctive. … It's also where there is more funding.” [SR]

He points out it can be difficult to get the initial funding since the funding can be reliant upon the ability to support larger projects, while the ability to support larger projects is reliant upon the funding.

> I think there’s a little bit of a weird chicken and egg thing going on. People say, “We want to do large projects like X at our institution,” but X wasn’t done by an institution, it was done with grant funding. I think every big project that you point to is a product of grant funding, and so to say we need to try to make it possible for everybody in every institution to do that kind of work is wrong headed. No institution can do that kind of work, no institution should be aspiring to do that kind of work without grant funding. [SR]

He also warns labs to think ahead, since the initial funding is not the end of the line. “Everybody who creates a grant funded digital project now has to talk about how they’re going to sustain it, in one way or another, but I think we, again, we don’t have good solutions to that, just because of the amount of ongoing maintenance that’s required in sustaining all the projects.” [SR]

3. SCALE

SR is emphatic that grant funding is not the only way to conduct research.

> I'm really reluctant that people approach digital scholarship through the prism of what can be done with grant funding. It's produced a lot of the backlash against digital humanities, because of the sense that it's inextricably tied up with external funding. It simply doesn't have to be. It's a matter of defining projects that can be done with the resources that are available to you.
> …
> All of the big, collaborative, digital projects have grant funding. That's what allows them to be big, collaborative projects. And yet, I strongly feel we give people the wrong impression of what digital scholarship is, if that is what we hold up as digital scholarship.
> …
> You're often asked by people at smaller institutions, I'm the only person doing it, how can I do a X project that costs half a million dollars involving eight people. The answer is simply, you can't, and you shouldn't be trying to. That doesn't mean that you can't do digital humanities scholarship, it means that you do it in a different way.
> …
> ***We need to really talk to people more about the different scales of digital scholarship that are possible***. [SR]

SR goes on to describe 3 ways for institutions to up-scale their projects without additional funding.

1. The digital domain itself and digital tools inherently allow for a much larger scale of data to be processed. "We don't require scale of other kind of scholarship done with other methods. The digital gives us possibility for scale." [SR]
2. Finding or developing in-house expertise. "Obviously if you've got somebody with a level of digital expertise, you have access to in your institutional library, you can upscale the work that you're doing." Additionally, "You can, in fact, develop the expertise to work as an individual in the same what that you would any other kind of project." [SR]

3. Collaborating with other institutions and experts. “You can enhance that, as you would on an analog project as well, by consulting with archivist, and librarians, and other people with expertise in the material that you want to work with, and the approach you want to take. That’s only one way of doing the bigger collaborative projects.” [SR]

While he encourages a balanced approach, he does concede that grant money makes things easier and points out that anyone can get it if they are motivated. “The goal should be producing work at the scale that the institution can support, and that the individual scholar can do. Now, that could mean going after grant funding. Grant funding is not only given to big institutions.” [SR]

4. TOOLS

Perhaps surprisingly, tools were not given a very high pedestal in terms of overall importance for DH scholarship. In fact, quite the opposite. Certainly, specific tools were mentioned, but usually as a means to an end, not as an end in and of themselves.

> If you are setting out to develop a digital humanities curriculum, or a digital humanities syllabus, or digital humanities Lab… I want to push back very strongly against the idea that you should start thinking about it in terms of the tools and skills perspective. Because, based on my own experience, giving job interviews, and the day-to-day sort of thing about what I do as a digital humanist, tools and skills are sort of the last things that come up. Really it’s more about a general philosophy or ethos. [BW]

Instead, the focus was consistently placed on the quality and value of the scholarly work and the metadata (the metadata often taking the form of XML).

> Somewhere along the way you're going to learn something about XML. And once you've grown in that community, it's all about the communication of, 'well, the reason we use XML is because, when everything else is gone you at least have the XML files that can be remediated into something else.' So, at least at the end of the day, no matter how the technology changes at least you have these files that are the scholarly pursuit and are the archives. If there's anything you're going to keep, it's going to be those files. [Everything else] is just going to break over time. [GG]

Even with DH scholars who try to develop their own software tools, it is rare for someone to develop something that is not already out there.

> Unfortunately, most of the people that are trying to come out with new models and new access policies and methods for digital resources, tend to come up with plans for new kinds of infrastructure. If you are really going to press them on explaining what they mean by the new infrastructure, they don't necessarily come up with anything that's not already out there. This is not a blind spot, it's an over-exposed spot, almost, in usual strategies to opening up digital resources for researchers and academics. [JZ]

Administrators and policy makers should also keep in mind from the previous section (discussing technical wherewithal) that the target skill range for most tools would be individuals with moderate to low digital fluency. So, breaking the bank to buy extremely complex and advanced tools only makes sense if there exists expertise to use it, as well as plans for a core capability to take advantage of the tools.

> In the sciences, labs do bet the farm on an instrument sometimes. You know, they go all in on some instrument. Here's an example: the biochemistry and biophysics department recently bought a new kind of 3D electron microscope and there's supposedly three ways of doing some kind of crystallography. It's a way of modeling a molecule, basically, for certain kinds of attributes. And for whatever reason this group in the UNC biochemistry department decided they were going to build what they call, a core capability around this really expensive instrument. So, they're basically saying, "We are going to build a research program around whatever we can process and get out of this instrument." And their advantage on other labs is that they could buy that instrument. They have access to it and they can get the data off of it. So, that's what makes them competitive. [DB]

An equivalent in the DH domain might be for a DH lab to buy highly specialized text analysis software or access to AI technology. DB states regarding the choice of

investment in tools and technology, “So, the labs are specialized, not generalized. … and they either get it right or wrong.”

5. THE CHALLENGES AND VALUE OF MAINTAINING DIGITAL COLLECTIONS

As was mentioned earlier in the section about funding, obtaining grant money is just the beginning of the journey. Maintaining a digital collection can be costly and time consuming, but it can also be the primary source of academic value for an institution.

> Maybe this is pushing things too much, but when I think about how money works in the humanities, I don’t always know how much databases are recognized, as they should be, for creating knowledge that can be used by other people.
>
> …
>
> The question is do libraries having the space and willingness to want to house this [digital material] perhaps for an indefinite amount of time? [ML]

SR states the details even more clearly:

> The more projects we develop, the more resources it takes to sustain them, and we're not getting any more resources to do it. A library would have exactly the same challenges that we do, in that no one is directly funding them to sustain resources, even as everybody is pointing at them to do it.
>
> …
>
> I say that as somebody with 20 years-worth of projects that we struggle to keep online here at the center, which is an activity that isn't funded by grants. We have to squeeze out money in other ways to do it. [SR]

Maintaining a digital collection can cost more than just money, depending on how an institution decides to host it. If they decide to host everything in-house, they need to worry about all of the infrastructure, maintenance, and security concerns that go along with it.

> The problem is there's so many legacy server environments that ran that have fundamentally bad security flaws, where it becomes a question of… well, you could put it on a server and people could access it but that opens up the server itself to other attacks. [GG]

In house hosting solutions are probably not an option for all but the wealthiest of libraries.

> Very, very wealthy libraries with a significant IT staff have some chance of doing that, but it would take all their time. That's not what their job is. Data sets I think are kind of easy, other kinds of digital projects are much, much more challenging, and we have lost a whole bunch of them. [SR]

However, with proper implementation, solutions for this problem are available.

> Any infrastructure that we need in terms of software tools is already there. Any infrastructure that we need on the hardware side is quite easily bought actually. It's all standard equipment, standard servers, standard server software. Any digital infrastructure you need is basically already out there. It's a matter of estimating how much you need of it and how you can acquire it most cost efficiently. [JZ]

This is not an issue that decision makers can or should avoid, for two reasons. First, because the DH scholars are already in need of a solution to this issue. Any institution intending to provide support and services for DH scholars MUST address this issue.

> It's always a challenge to figure out "Well, okay. We want to make this online archive. If it's not going to be hosted by the library, because it's our derivative archive? Should it be hosted by our co-sciences? Is it the responsibility of the department?" What we've also ended up doing is getting a Blue Host or Dream Host or something, a host account, completely outside of any University system, ***and created things ourselves there, because there's no internal support for creating a digital archival exhibition that is outside the realm of library prioritized projects.*** Even within the college, a few things might rise to the level of centralized support. [VS]

DH scholars should NOT be forced to create their own solutions to this issue, because inevitably the dataset will just be lost, which is tragic, considering the second reason why institutions should address this issue: ***digital collections are some of the most valuable assets a digital lab has to offer.*** SR states, "We have 20-year-old projects still attracting millions of users." VS explains it this way:

> We have a few big collections like the history of advertising collection, for example, that we know a lot of people are very interested in and they prioritize putting those materials online and then work with faculty and other researchers who want to do something with that collection. So, that becomes a feature.
> …
> ***I think the exhibition objectives are partly to raise awareness for the library and the resources that they have and hopefully that translates into greater funding and donations and all those kinds of good things.*** [VS]

Various institutions are starting to recognize the value of these assets and are beginning to provide options for hosting digital collections. "Social sciences journals also take some responsibility for collecting and archiving data sets. Humanities journals are overdue to start thinking about that." [SR]

It's certainly an investment, but just like an analog library or museum, the true value is in the collections.

> I think that is the much smarter move for digital humanities activities. But it's just my opinion. It's not trying to produce a general solution that will satisfy whatever project. What you're doing is you're cultivating a resource for study. And so you're maximizing the salience or tractability of what somebody, who's informed about the discipline, or they're just a true believer about the significance of some object or set of objects that's going to matter. I think that that's a good way to go. Because basically you're making that collection, that cultivated collection, an asset. [DB]

Is an asset that attracts millions of users worth the money? In some cases, that's quite literally the million-dollar question. It's something each institution must decide for themselves. If an institution does decide that there is value in creating and maintaining these digital collections, administrators should consider putting in places policies that require the collections to uphold certain standards, this will only further increase the value of the datasets and provide a threshold of entry, eliminating poorly constructed datasets. VS describes this process, "In the cases of doing online project where we do

want to have some things that exist longer, then we try to adhere to standards like Dublin Core using tools like Omeka." BW adds, "They give really detailed workflows and tutorials about how people can actually use the data that's there. So that it's not just a matter of putting the data online, but actually helping scholars see the steps that are necessary to begin processing it on their own."

Incentives

As the literature review pointed out, the environment a DH "ecosystem" needs to thrive is intensely focused on building bridges between specialties and rewarding all forms of labor in proportion with their effort and value. The interviews reflected this theme and provided specific examples of the need to make sure that everyone has appropriate incentives to participate.

> The humanities needs to get into computer science departments and figure out, what can we bring? If the computer science department is weak, you throw the computer science department in with the humanities and you start to build tools. That's a model that works, but at UNC we have a very strong computer science department and they do image processing. So, maybe we need to think, as humanity scholars, what interesting questions can we bring to them about images that would challenge them and provide some interesting interactions? [GG]

GG's desire to bring something interesting to the table may raise doubt in some minds as to what the humanities might be able to contribute to a more technical field. However, SR lays those doubts to rest, pointing out that the humanities brings funding of their own that benefits everyone. "We work a lot with colleagues in the history department who provide expertise for various projects we do. Sometimes we work with them to get funding for their own projects." [SR]

He goes on to point out that the lab he works in is not part of a technical department or even free floating. “We're formally part of the history department, which I think people lose track of. They think we're somehow free floating in Mason. We're not. We get a lot of our institutional support filtered through the history department, rather than doing it from other parts of George Mason. That includes the leadership of the center.” [SR]

Most importantly however, the humanities can also give awards and recognition which can directly influence funding as well.

> Digital Harlem, the website, won first AHA prize for innovation in digital history, and a prize from the American Library Association. We were able to use that to get further funding from the Australian Research Council. They rejected our proposal for the Riot project the first time we put it in. Without those prizes - on the back of some reviewers who said we weren't doing anything innovative or interesting on the digital mapping - when we won the two prizes that were contradicting that reviewer, we put it back in and got the money. [SR]

Lest someone think that the digital lab in question here is a small digital lab owned by the history dept, it’s not…

> We're pretty much the oldest digital humanities center in the United States, some people at UVA would quibble about that, but we're certainly the most high profile one of research centers in the world that has a regional and international reach. That's the product of 20 years’ worth of people and projects, it's the product of Roy Rosenzweig's original vision of what was going to be possible in the digital realm. [SR]

The point being, humanities has a lot to offer the digital/technical side. This can be and has proven to be a very healthy and symbiotic relationship/collaboration in many cases. However, the current model for what deserves recognition and scholarly awards needs to change within to humanities to support this new collaborative environment.

ATTRIBUTION & INCENTIVES NARRATIVE:

JZ states things very clearly:

> It seems that research incentives and research structures within the institutions are still very much in line with how we did research, basically 20, 30, 40 years ago. As long as that doesn't change, we're not moving quickly or very efficiently in a situation where the connection between researchers, librarians and digital resources is going to be this tremendous explosion of computational research. [JZ]

VS describes the current situation of what humanities tend to recognize as scholastic work. “It's very much still a written publication type of field in terms of what they count as scholarship in general. Coming out of English Literary Studies that's still true. It's changing at some institutions and our University is starting to look at broadening its understanding as what constitutes scholarship, but it's a pretty recent development.” [VS] This needs to change, otherwise no one is incentivized to do this work.

> Just in terms of finding materials and then in terms of making them available online, or programming, or analyzing them, or whatever… [faculty members] tend to want to outsource that to somebody, whether that is a grad student or a librarian or some other type of assistant or collaborator. We do already see some hierarchy happening around that. A privileging of certain kinds of work that would be retained by the faculty member versus delegated to somebody else based on what's likely to get them credit. [VS]

This is not to blame the faculty member, as VS also points out. “I think that for faculty members who are not yet tenured, it could be a career killer to spend too much time doing this kind of work.” [VS]

Right now, people who work at digital labs are perceived to be in a “support” role, when in reality, they should be perceived to be in a scholastic role. VS outlines the struggles that occur due to these roles being viewed as “support.”

> They do promote the idea of librarians being collaborators, and librarians want to be collaborators. I think it's just the counter message that is received by what's institutionally operational makes it feel like there's less opportunity to actually fulfill that collaborator role in the way that people sometimes want to do. There's discussions about doing it. It's just how does that translate? [VS]

As discussed earlier, the awards and recognition that the humanities can provide is critical. This scholastic recognition would not only "allow" at an institutional level, but it would motivate people at an individual level, and perhaps ignite the "explosion of computational research" that Joris spoke of. As it is, the librarians and other "support" people are trapped in a catch-22, where more time is needed in order to do things that are institutionally considered scholarly, but at the same time, pressure at an institutional level prevents them from investing more time.

> They are allowed to do [this work] in a service provider role, but when I've tried to buy out time for people to work on projects in greater depth, I've been told we can't do that. I know that's a source of frustration all around that if you want to do these deeper interventions, for the most part, the people who do have subject area expertise in metadata and things like that are not able to spend the time, because it's perceived as spending too much time on one thing as opposed to participating in a broader more systemic support role. [VS]

There is little opportunity to earn scholastic recognition without investing quality time, but scholastic merit should not be measured by time invested. Instead it should be measure by value added to the project. "I think the best kinds of relationships are those that are truly collaborative and that truly acknowledge the scholarly work of everyone involved in the process. [A process] that engaged the librarians and staff involved as partners in the research, not simply service providers." [BW]

Despite the lack of institutional support, at an individual level, there is a conscious effort to recognize the help that the librarians and digital labs folks provide as scholastic.

> I think there's a point where the librarian provides enough material to be a second author on a paper.
> …
> There's a librarian at Duke that I partner with, and I'm doing a project on the HB2 tweets that happened last year. And, he was able to use the software they have at Duke for social feed managers to get more of the tweets from HB2… So, he provided much more information than I could ever find on my own, and that information is pretty much this whole project.
> …
> He and I have talked…*He'll be a second author on that paper*, because the method of programming something in Twitter to scrape all those tweets… That is such a huge part of the project that *I just don't feel like it's right not to include him in the paper.* [GG]

BW claims: "We have a suggested citation that includes, actually, the whole praxis program team, so that the contributions of everyone involved are acknowledged." [BW] DB states, "If it were me, I would… Whoever it is that went and did the presenting of Kant's complete works in the format that they did on Korpora.org: that deserves a PhD in philosophy or history or something like that. I don't think there's any question." [DB]

But where should that line be drawn? "I think we're still trying to figure out like what the multi-authored paper really looks like." [GG] Even if institutions were to agree that the boundaries need to shift, where is the new boundary? DB provides some very good measurements regarding what should be considered scholastic work:

> Somebody would have to know an awful lot of things about the very specific domain area in order to do [Kant's complete works] competently. You know, it's an astonishing accomplishment! And one of the ways that I look at it is, I think *the contribution that cannot be accomplished by a machine, that a computer itself cannot do, I think that that's part of what provides its intellectual merit.* [DB]

In other words, if something is "plug-and-play" it doesn't rise to the level of scholastic merit. However, if it requires an understanding of the specific domain area,

then it should count as scholastic even if it is the act of creating and curating a digital collection. GG adds to this:

> I think that hidden labor of curating a collection, or creating a catalog record of a book, or the descriptive bibliography of what a book is about is extremely scholarly. And there's been notions of power and authority built into the academic structure that, for a long time, have wiped away the immense labor of the librarian.
>
> What does it really mean to do humanities scholarship and what are the humanities methods? Especially when we're trying to figure out what's the humanistic approach to running an algorithm on a bunch of text…
>
> How is that different then computer scientist, or someone from statistics? I think these questions are starting to get asked. We've had to reevaluate all of our old notions because were moving into a different form. [GG]

GG goes on to define additional boundaries as to what does and does not deserve merit as scholastic work. Specifically, analyzing and marking up text with meta data is scholarly, while decorating the text with fancy website designs is not scholarly.

> The curation of the text, and marking that up, is the scholarly pursuit. The user interface and flashiness of putting in a particular edition of the website is nice, and I think there's a place for it in terms of engaging the public and engaging the wider audience, but at the end of the day, is this data set good enough? And is the data set really a scholarly pursuit? Or is it building and designing a website? [GG]

Basically, to really do the research that needs to be done, scholars have to know the text intimately and have to build or rebuild the collections. In fact, the librarians have always been a part of the research, but that labor has been hidden because it was just called library work or archiving. For example, when a collection is built, things are cataloged, tagged, and interpretation of the material has already started to take place by stating, this goes with that, and this is associated with that, and by grouping things together and making them searchable.

A scholar must know the material in order to do that. So, the question arises, to what degree should this hidden labor be given scholastic recognition? Digital humanities is simply revealing this hidden scholarly labor that librarians and archivists have always done, but now it is more visible because it's not being done as fast as scholars would like at this point.

VS comments on the irony of how an article about a collection is considered more scholarly than the act of creating the collection, even though the act of creating the collection is so much harder:

> In terms of research projects, if I write an article based on the work, describing the work, then that might count more as scholarship or doing an exhibition based on the work. The final results would be something that would be more recognized as scholarly work, but creating metadata standards and creating databases and things like that are much harder. [VS]

JZ also supports this viewpoint. “I think the one obstacle has been the ongoing unwillingness to recognize the creation of digital collections as a form of scholarship and interpretation.” [JZ] He further elaborates that this work is more visible to digital scholars because they are often doing similar work to librarians. “Digital scholars are often aware of this, because they're doing parallel work to what librarians do, and they want to be credited as doing scholarship when they do it. That goes with recognizing that that's what librarians are doing as well.’” [JZ]

ML describes an increasing necessity for this sort of recognition, and the ease with which the digital element makes this tracking of who did what.

> I get the sense that there's a growing recognition and conversations about this hidden labor and how to make sure that it's legible, because even in the Blake archive,

we're quite careful about documenting who's changing what, which assistant, when we're making changes to the files or the image mark up, you can click through and see who is responsible, which is also nice as a graduate student because then it's not just that my work is being completely effaced and not showing up anywhere. [ML]

I have hope for the near future that a focus on book history and publications is going to bring with it a legitimization of this work that librarians do, and legitimizing the work that they're doing as scholarship. And when that comes, I think that's going to help a lot of things. [GG]

Conclusion

The introduction to this paper pointed out that one of the most significant common denominators in user needs among DH scholars is the need to enable collaboration across highly dissimilar academic disciplines and professional skill sets. In order to better unpack the concept of collaboration, a market analysis approach was chosen, and a model was introduced that defined the groups involved in DH collaborations and their relationships to one another. The literature review endeavored to describe the resources and needs, cultural norms, and impediments and dysfunctions each group in the model brings to any given collaborative attempt. Finally, the analysis section provided insights on the DH collaborative market from the perspective of humanities scholars through a series of semi-structured interviews with seven well-established DH scholars. The interview data was analyzed using a grounded-theory, qualitative approach.

The primary research question driving this study was a desire to understand from the perspective of scholars in the humanities, what the barriers are to the adoption and growth of digital methods and tools. The goal was to contribute usefully to the decision-making process at the level of university administrations, library boards, and other funding organizations with a stake in the outcome of DH scholarship. Any of these organizations that have invested in the significant cost and effort of digitizing their collections, particularly any archival collections unique to their institution, should have a profound interest in encouraging the effective and scholarly use of them. Specifically,

they should be interested in whether the digital collections enable new and innovative forms of humanities research and produce new methods and approaches to analysis of the collections.

The interviews provided three key insights in answer to the overall research question:

1. Metadata is data, and the curation of such data is a scholarly pursuit
2. Computational software and tools do not substitute for human collaboration
3. Hidden labor is lethal to collaboration; incentives should be shared equally among the three major players in DH – data science experts, information seeking experts, and humanities experts

The data the interviews provided on the relationship between humanities scholars and archivists/database developers showed that *it is absolutely critical to the continued growth of DH for the data production and scrubbing process to be considered scholarly.* The process of converting documents to data is time consuming and laborious, can account for a significant portion of the time and effort invested in any given DH project, and often requires the input of multiple individuals from different areas of expertise. If none of that labor counts as scholarship, it profoundly undermines the ability of anyone to engage in the process without access to unrealistic amounts of independent funding.

Interviewees indicated that *access to consultative resources were critical to facilitating efficient access to and use of digital collections, as well as the use of any advanced analysis tools or software offered in a "digital lab" context*. Institutions should consider adopting the "4th Librarian" with "just for me" services, aimed at the "middle" group of users with a "fluency" in technical skills. It makes sense to limit the time and resources of the lab to a motivated minority of individuals who have demonstrated

sufficient desire to learn a basic competency and that come into the lab with at least a basic understanding of programming. The consultants or "4th Librarians" can then focus on providing creative solutions that allow scholars to break through or at least go around accessibility walls to efficiently achieve the skills and materials they need to accomplish their research goals without trying to achieve a level of hyper-expertise that could actually be detrimental to their career.

Finally, interviewees indicated that administrators need to make a greater effort to acknowledge and reward all labor involved in collaborative DH projects. The widespread issues of hidden labor, "digital miracle workers," undefined responsibilities, and lack of credit for value contributed stunt the long-term viability of DH collaborations if they don't prevent them from ever occurring in the first place. Institutional leadership should be very careful in how it defines success. Otherwise, librarians, technical consultants, archivists, and other "support" people are trapped in a catch-22, where more time is needed in order to do things that are institutionally considered scholarly, but at the same time, pressure at an institutional level prevents them from investing more time. In addition, institutions should be careful not to place so much value on completed projects that can be published or exhibited that they lose sight of the tremendous long-term scholarly value of the data itself. The interviewees emphasized that the *digital collections themselves are some of the most valuable assets a digital lab has to offer, and offering the ability to build on previous scholarly effort is the path to DH success.*

This study limited its scope to the perspective of DH scholars and their views of the barriers and relationships within the DH market. Future studies could expand upon this by conducting similar interviews with the other members of the DH market. Ultimately a

complete picture might emerge showing the overlaps and gaps in perspectives between the various groups. This would provide institutional leadership with key insights on how to remove barriers and build incentives to increase involvement and progress in scholarship in the Digital Humanities.

Bibliography

Anderson, Sheila. "What Are Research Infrastructures?" *International Journal of Humanities and Arts Computing*, vol. 7, no. 1-2, 2013, pp. 4–23., doi:10.3366/ijhac.2013.0078.

Bales, Andrew. "American Lynching: Uncovering a Cultural Narrative." *American Lynching*, University of Cincinnati, 2016, www.americanlynchingdata.com/index.html.

Benardou, Agiatis, et al., editors. "Introduction: a Critique of Digital Practices and Research Infrastructures." *Cultural Heritage Infrastructures in Digital Humanities*, Routledge, 2018.

Blouin, Francis X., and William G. Rosenberg. *Processing the Past: Contesting Authority in History and the Archives*. Oxford University Press, 2013.

Boyles, Christina, et al. "Precarious Labor in the Digital Humanities." *Digital Humanities 2018: Book of Abstracts*, edited by Jonathan Girón Palau and Isabel Galina Russell, Alliance of Digital Humanities Organizations, Mexico City, 2018, pp. 47–52.

Burdick, Anne, et al. *Digital_Humanities*. MIT Press, 2012, *MIT Press*, mitpress.mit.edu/books/digitalhumanities.

Chappelow, Jim. "Principal-Agent Problem Definition." *Investopedia*, Investopedia, 10 Sept. 2019, www.investopedia.com/terms/p/principal-agent-problem.asp.

Charmaz, Kathy. *Constructing Grounded Theory*. 2nd ed., SAGE, 2014.

Cordell, Ryan, "How Not to Teach Digital Humanities." *Debates in the Digital Humanities 2016*, 1st ed., ser. 2, University of Minnesota Press, 2016. 2.

Derrida, Jacques, and Eric Prenowitz. "Archive Fever: A Freudian Impression." *Diacritics*, vol. 25, no. 2, 1995, pp. 9–63., doi:10.2307/465144.

Edel, Abraham. "More on Knowledge and the Humanities." *Journal of Aesthetic Education*, vol. 21, no. 4, 1987, pp. 95–101., doi:10.2307/3332834.

Encyclopaedia Britannica, "*Market*." Edited by Joan Violet Robinson, Encyclopaedia Britannica, Inc., 10 Feb. 2017, www.britannica.com/topic/market.

Hoekstra, Rik, et al. “Data Scopes: Towards Transparent Data Research in Digital Humanities.” *Digital Humanities 2018: Book of Abstracts* , edited by Jonathan Girón Palau and Isabel Galina Russell, Alliance of Digital Humanities Organizations, Mexico City, 2018, pp. 210–212.

Hollingsworth, Rogers. “The Snare of Specialization.” *Bulletin of the Atomic Scientists*, vol. 40, no. 6, 1984, pp. 34–37., doi:10.1080/00963402.1984.11459243.

Kuhlthau, C.C., and S.l. Tama. “Information Search Process of Lawyers: a Call for ’Just for Me’ Information Services.” *Journal of Documentation*, vol. 57, no. 1, 2001, pp. 25–43., doi:10.1108/eum0000000007076.

Langmead, Alison, et al. “Towards Interoperable Network Ontologies for the Digital Humanities.” *International Journal of Humanities and Arts Computing*, vol. 10, no. 1, 2016, pp. 22–35., doi:10.3366/ijhac.2016.0157.

Lu, Li, and Y. Connie Yuan. “Shall I Google It or Ask the Competent Villain down the Hall? The Moderating Role of Information Need in Information Source Selection.” *Journal of the American Society for Information Science and Technology*, vol. 62, no. 1, 16 Sept. 2010, pp. 133–145., doi:10.1002/asi.21449.

Malazita, James W. “Digital Humanities 2018.” Red De Humanidades Digitales A. C., *Digital Humanities 2018 Puentes-Bridges Book of Abstracts* , 2018, pp. 235–237.

Olson, Hope A. “The Power to Name: Representation in Library Catalogs.” *Signs: Journal of Women in Culture and Society*, vol. 26, no. 3, 2001, pp. 639–668., doi:10.1086/495624.

Padilla, Thomas. “Collections as Data .” Library of Congress, *The Signal*, Sept. 2016, digitalpreservation.gov/meetings/dcs16/tpadilla_OnaCollectionsasDataImperative_final.pdf.

Pascoe, William Dudley. “Rapid Bricolage Implementing Digital Humanities.” *Digital Humanities 2018: Book of Abstracts* , edited by Jonathan Girón Palau and Isabel Galina Russell, Alliance of Digital Humanities Organizations, Mexico City, 2018, pp. 469–471.

Rhody, Lisa, et al. “Expanding Communities of Practice: The Digital Humanities Research Institute Model.” *Digital Humanities 2018: Book of Abstracts* , edited by Jonathan Girón Palau and Isabel Galina Russell, Alliance of Digital Humanities Organizations, Mexico City, 2018, pp. 659–678.

Rosenzweig, Roy. “Collaboration and the Cyberinfrastructure: Academic Collaboration with Museums and Libraries in the Digital Era.” *First Monday*, vol. 12, no. 7, 2007, doi:10.5210/fm.v12i7.1926.

Speck, Reto, and Petra Links. “The Missing Voice: Archivists and Infrastructures for Humanities Research.” *International Journal of Humanities and Arts Computing*, vol. 7, no. 1-2, 2013, pp. 128–146., doi:10.3366/ijhac.2013.0085.

Thomasen, Ulf Porup, and Peter Kastberg. “Increasing Specialization of Knowledge Leads to Increasing Fragmentation.” *Journal of Organizational Knowledge Communication*, vol. 2, no. 1, 2015, pp. 1–5., doi:10.7146/jookc.v2i1.22793.

Unsworth 2006 Unsworth, John, et al. Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences. The ACLS Report. 2006. http://www.acls.org/cyberinfrastructure/ourculturalcommonwealth.pdf.

Van Zundert, Joris. “Barely Beyond the Book?” *Digital Scholarly Editing: Theories and Practices*, edited by Matthew James Driscoll and Elena Pierazzo, Open Book Publishers, 2016.

Van Zundert, Joris. “If You Build It, Will We Come? Large Scale Digital Infrastructures as a Dead End for Digital Humanities.” *Historical Social Research / Historische Sozialforschung*, vol. 37, no. 3 (141), Controversies around the Digital Humanities, 2012, pp. 165–186.