Investigating how South African Humanities Researchers Engage with Digital Archives

by

Khanyisa Mtombeni

MTMKHA001

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In partial fulfilment of the requirements for the degree

MASTERS IN INFORMATION TECHNOLOGY

Faculty of Science

University of Cape Town

Date:

Supervisor: Hussein Suleman

Declaration

I, Khanyisa Mtombeni, hereby declare that this dissertation titled “Investigating how South African Humanities Researchers Engage with Digital Archives” is original work carried out by me under the supervision and guidance of Hussein Suleman. This is the case throughout except where acknowledgements indicate otherwise. No part of this work has been, is being, or is to be submitted at another university for a fulfilment of another degree.

Signed by candidate

Signature: ……

Date: 11 August 2021 (revised)

Acknowledgements

I would like to appreciate my supervisor Hussein Suleman for providing guidance and feedback throughout this project and for his invaluable constructive criticism and advice in the course of its completion. The meetings and conversations we had were invaluable. I would also like to thank my family for their never-ending support and all the people who provided me with the facilities that I required and were conducive to me completing this research.

Abstract

OBJECTIVE: Despite technological developments in the Digital Humanities space, it is unclear that the facilities offered by digital archives support the needs of Humanities researchers in developing countries. The purpose of this thesis is to investigate how South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities.

METHODS: This thesis utilizes non-random convenience sampling. A feature determination study provided the sampling frame, defined the scope for the survey tool, and was used to uncover trends in digital archives development in South Africa. A self-administered online survey was conducted with Humanities researchers in South Africa to answer the research question. The thesis utilises basic descriptive statistics in its attempt to study and interpret the responses of participating researchers.

RESULTS: 102 participants responded to the online survey. Despite many South African digital archives having the functionality to discover, browse and search collections, they are missing the features for collaboration, accessing and managing resources. Only 20% of the survey respondents are satisfied with South African digital archives' process of making content easy to find and accessible, whereas 48% of the respondents consider themselves users of complex digital resources, 44% have the knowledge and experience for using Digital Humanities tools and services, and more than 70% find technology to be useful for learning and teaching.

CONCLUSIONS: The usage of archives and their functionalities vary widely. Users have stronger preferences for tools that support basic discovery and personal and collaborative research, but many consider existing support for basic features to be inadequate. In terms of advanced functionalities for managing digital resources, users are interested in these to varying levels, but the inadequate support means that these are still somewhat speculative.

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1. Background

Modern Humanities researchers such as historians, archaeologists, linguists, and individuals doing evidence-based exploration of questions or hypothesis that are important to the Humanities disciplines often engage with large collections of textual and physical material that is in digital format. These digital collections are ideally stored and managed in digital archives with a set of useful services provided to researchers. This has become standard practice all over the world as standard archiving tools have emerged. Digital archive tools such as Atom and Archivematica are popular in the Digital Humanities domain and, technically, function similarly to other archiving tools. However, digital archives, which are repositories for storing, managing and preserving digital content over long periods, are not necessarily easy to set up, use and manage. In low resource environments, as is the case with most African countries, these tools are also not widely adopted due to resource, technical and funding related issues (Suleman: 2019).

Among Humanities scholars, the use of Digital Humanities tools has remained marginal. According to Terras (2016), in 2005, an estimated six percent of scholars in Humanities went beyond using general-purpose information technologies and digital resources for complex research activities. By 2012, an estimated 10 per cent of Humanities scholars systematically used digital resources. In 2014, nearly 50% of faculty members in the Humanities had created or managed digital resources. A strong desire exists among tool developers to learn more about researchers' needs and practices with digital archives as the user base and user needs

continue to broaden and change. The change in user needs and expectations presents major considerations in the way archiving systems are designed and developed.

1.2. Problem Statement

Technological developments have made possible the creation of digital archives of scholarly knowledge. These developments have contributed to making the field of Humanities data-rich and have made access to tools and resources of knowledge discovery easier (Urberg: 2017). Despite these developments, developing countries like South Africa still face challenges in creating and maintaining digital archives. Also, very few studies explore how Humanities scholars integrate information technologies into daily research practices (Given and Wilson: 2018). Because of these factors, it is not clear how Humanities researchers use digital archives and tools in their research, teaching and other academic activities. It is also unclear that the facilities offered by digital archives meet the needs of Humanities researchers in South Africa.

1.3. Research Question

How do South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities?

1.4. Methodology

This thesis will begin with a feature determination study of existing archives and archiving tools to determine the extent of features available to users. Then, a survey was conducted with Humanities researchers in South Africa to answer the research question in the context

of the technical features of known systems. This was a self-administered survey completed via the Internet. The thesis utilises quantitative methods in the form of basic descriptive statistics in its attempt to study and interpret the responses of participating researchers. It employs a non-random convenience sampling approach to select Humanities researchers for involvement in the study.

1.5. Rationale and Significance of Research

There is value in studying existing digital archives and archiving tools to understand the extent of features available to users. This provides an opportunity to engage with users to understand how they interact with current digital archives and tools. The information gathered from users will aid the development of tools that are potentially more effective, appropriate for the South African environment and widely used by South African researchers.

1.6. Structure of the Dissertation

The first chapter provides a summary of the research problem and explains the purpose and significance of the research and how the study will be carried out. The second chapter is a literature review, which gives a historical development of Digital Humanities and digital archives, identifies the research that already has been completed in the area of Digital Humanities and archiving, and provides an analysis of the current information relevant to the research topic. The third chapter is the methodology chapter. It explains the methods used to gather information and data to answer the research question. The fourth chapter provides analysis and discusses the results. Chapter five briefly revisits the rationale for the study, presents a summary of the main findings and gives recommendations for future research.

Chapter 2: Literature Review

2.1. Introduction

Digital Humanities research and teaching occur at the crossroads of digital technologies and Humanities. As stated by Terras (2011), Digital Humanities aim is to create and use applications and models that allow for new kinds of teaching and research. For this to be accomplished, Humanities, Computer Science and related technologies must converge. Digital Humanities studies the impact of research and teaching techniques on cultural heritage, heritage institutions, libraries, archives and digital culture. Meeks and Grossner (2012) describe Digital Humanities as Humanities research using digital objects, tools and methods. It results in publications and the development, extension and annotation of digital archives. It also involves the development and improvement of digital tools.

Developing countries, including South Africa, have taken interest in digital archives and the preservation of digital records, but several challenges stand in the way (Yadav: 2016). Despite the design, creation and maintenance challenges faced by digital archives, digital materials are not going away and there will not be any less in the future than there exists today (Urberg: 2017). It is therefore critical that serious thought is put into the development of digital archives which incorporate user experience, engage with collections at various levels and perspectives, and that allow researchers to compile and analyse materials whilst staying fit for the needs of researchers in developing countries (Janco: 2014).

As a step towards investigating how South African Humanities researchers engage with digital archives, this chapter gives a historical development of Digital Humanities and digital archives, identifies research that has already been completed in these areas and provides an analysis of the current information relevant to the research topic. The chapter also investigates methods and tools necessary for building digital archives that meet the needs of researchers, including the challenges facing digital preservation in developing countries.

2.2. The Context for Digital Archives in Humanities

Busa (1980) traces computing in Humanities back to Father Roberto Busa, an Italian Jesuit priest who, in 1949, took on the demanding task of making an index of all words in the works of St Thomas Aquinas and related authors using a concordance programme. According to Ries (2019), the early beginnings of archiving practice are mostly associated with archivists, librarians, archival and information scientists, and humanists. He asserts, “Digital archives… are rich, diverse and multi-faceted treasure troves for historians, political scientists, sociologists, philologists, literary scholars, art historians, digital humanists, and researchers from other Humanities disciplines”. Baucom (2019) believes that the exponential increase in the use of personal computers, the arrival of the Internet in the 1990s, and the explosion in the number and types of digital objects has led towards a comprehensive and collaborative approach to digital preservation.

Digital computer technology continues to be of relevance to the various fields of Humanities (Saltz: 1997). An example of where digital computer technology and digital archiving continues to be relevant is in the fields of Performance Art, Art History and Archaeology. Sant

(2014) says that digital technology has been able to resolve creative, practical and scholarly challenges within the field of Performance Arts, despite performance arts not being a textual, but rather, visual, auditory and physical field (Sant: 2014). Art History uses computers to order and analyse data about artworks (Raben 1991). According to Eiteljorg (2001), Archaeology looks at artefacts, buildings, organic material and landscapes and takes into consideration the context of the objects excavated, as a result, good record keeping is important to the field as records preserve context, artefacts and other aspects of the excavation.

Computing has also contributed to the fields of Linguistics, Literary Studies and Musicology. Linguistics, which is the study of human language, looks at the structure and usage of both spoken and written language in relation to sciences such as Psycholinguistics and Sociolinguistics, and, in modern times, Computer Science. According to Zerkina & Lomakina (2017), computing contributed significantly to the development of Linguistics, and yet, without linguistic discoveries, there would not be a computer and computer programs. The digital archiving of linguistic work preserves digital language documentation and makes it reusable and accessible to other linguists. In Literary Studies, which is the understanding and preservation of literary texts, according to Trimpi (1970), three requirements have been vital for getting literary computing recognized by mainstream scholarship. The three requirements include (1) access to high-quality electronic text, (2) sophisticated software that lets the user dictate the terms of analysis, rather than the software defining the terms and (3) unlimited computing power and storage. Based on Byrd and Crawford (2001), as more music becomes available on the Internet, so is the need to have methods of searching for required music information, and as interest in digital archives continues to increase, so will the interest in

preservation and retrieval of music. Hurron (1999) believes that as more music data becomes available, the music research methodology and the discipline of Musicology will become richer in data.

The interdisciplinary nature of Digital Humanities and the evolution of Computing has significantly contributed to new approaches for archiving, research, learning, teaching and collaboration. Studies in Digital Humanities continue to emphasize the important role of computing in Digital Humanities and the need for sophisticated analytical technologies for search and retrieval. There continue to be important digitization issues that need to be resolved in computing by Humanists building digital tools for research and learning. The issues include what to digitize, how to represent digitized objects, how to digitize objects accurately, and how to use digital tools to enhance research and learning activities (Rockwell and Mactavish: 2004).

2.3. Models for Building Digital Archives

Several models and theories of digital archives and digital libraries have been developed to guide digital archives and Digital Library developers towards building tools that will meet the information needs of various users. The Delos Reference Model and the 5S Theory Model are abstract and prominent models for building digital libraries (Isa et al.:2013). The 5S framework defines the basic set of features for a Digital Library as seen in Figure 1 (Murthy et al.: 2007). The DELOS Reference Model captures the entities and relationships of the Digital Library universe as shown in Figure 2 (Candela et al. 2007). Nevertheless, there is no universal

reference model that covers all aspects of digital libraries and digital archives precisely and rigorously (Murthy, U. et al.:2010).

Digital archives and digital libraries are complex systems due to their interdisciplinary nature. An integrative theory of digital libraries is therefore necessary. According to Shen et al. (2008), The 5S framework is an integrative theory that is useful for defining Digital libraries. It defines the minimum set of concepts required for a system to be considered a Digital Library. The 5Ss refer to streams, structures, spaces, scenarios, and societies. The 5Ss can be used to define other Digital Library concepts including digital objects, metadata, collections, repositories, and services. According to Murthy et al. (2010), the 5S framework can be extended to provide support for complex objects, retrieval services, and superimposed information and services.

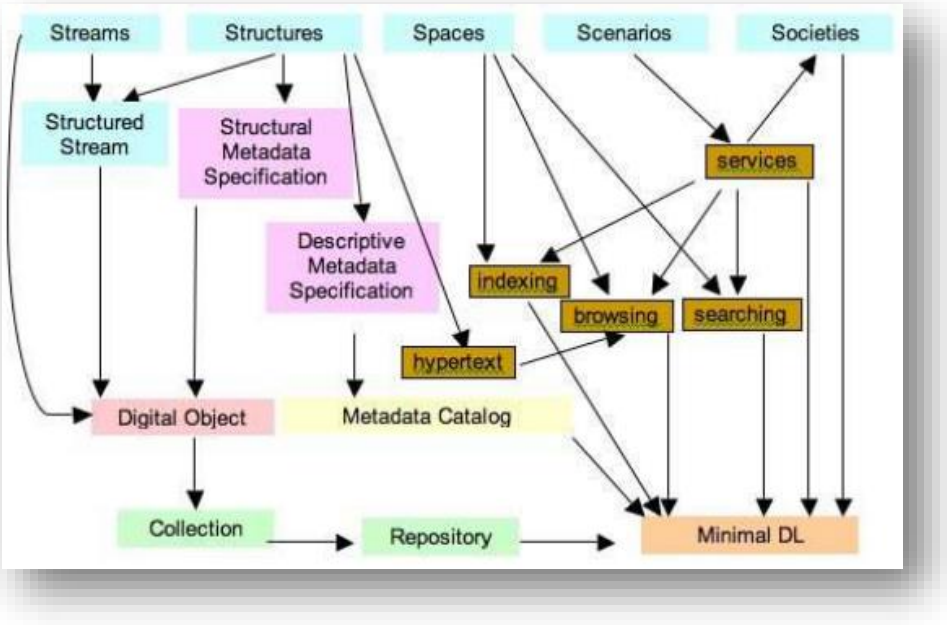

Figure 1: A minimal DL in the 5S framework (Murthy et al.: 2007)

DELOS envisions Digital Library systems with no borders or barriers to information. It envisions systems that are accessible everywhere and all the time and that offer a user friendly, dynamic, effective, efficient, and interactive environment. The envisioned system is human-centric, communication-centric and collaboration-centric (Ioannidis, Y. et al.: 2008). The Delos reference Model presents three ideas of 'systems' which play a central and distinct role in the Digital Library development process: Digital Library, Digital Library System, and Digital Library Management System. Six core concepts influence and provide a foundation for digital Libraries: Content, User, Functionality, Quality, Policy and Architecture (Candela et al. 2007).

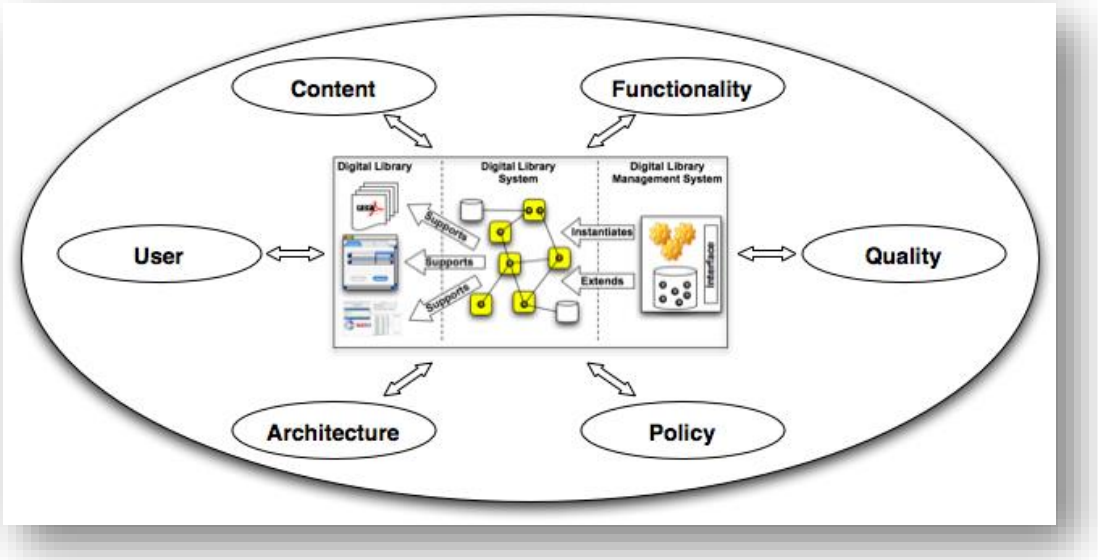

Figure 2: The DELOS Digital Library Reference Model (Candela et al.: 2007)

The Consultative Committee for Space Data Systems (2012) present a standard abstract model, called the OAIS (Open Archival Information System) reference model. The OAIS model is an interpretive and flexible model applicable in building digital archives for Humanities and social sciences. As depicted in Figure 3, the functional model describes the processes: Ingest, Data Management, Archival Storage, Preservation Planning and Access. These processes

describe the route information packages take through the archive and the activities and tasks that can be fulfilled by either a human or a computing system. The OAIS functional model communicates the division of responsibilities between the participants outside of the archive – users and producers of information – and the entity managing the archive. A study by Habert and Huc (2010) suggest the OAIS is ideal for the development of digital archives given the constant change and sometimes disappearance of digital technologies, be it software or hardware.

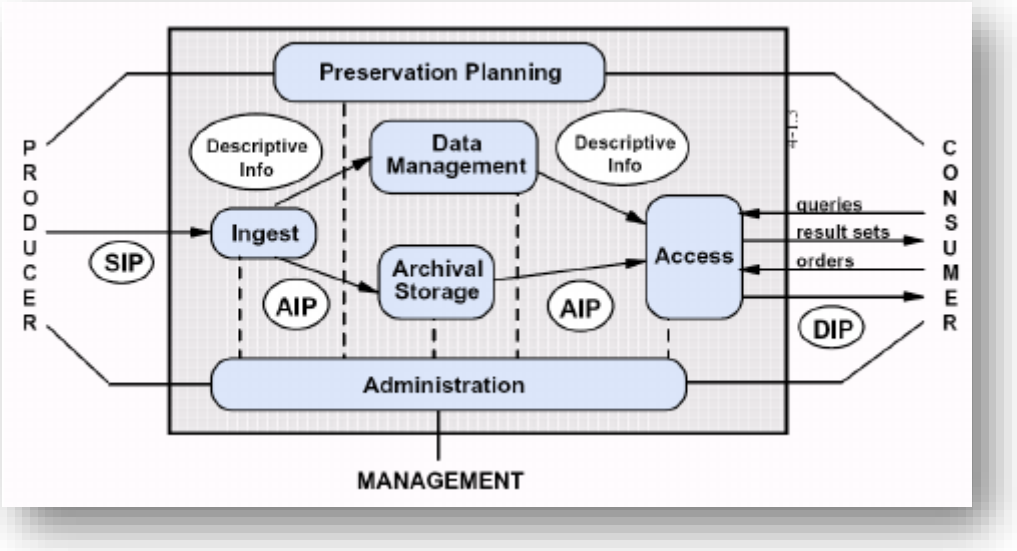

Figure 3: Reference Model for an Open Archival Information System (The Consultative Committee for Space Data Systems: 2012)

These three models provide a basis from which developers of digital archives can start thinking about building digital libraries and archives that serve the needs of African Humanities researchers. They create a path for the development process since programmers do not need to reinvent the wheel each time they develop a new application.

2.4. Methods and Tools for Research using Digital Archives

To bring efficiency to research using digital archives, Janco (2014) proposes the idea of a unified platform that allows researchers to search, compile and analyse materials from digital archives. This platform, according to Janco (2014), would facilitate research efforts by utilizing images, scans, spatial and GIS data, text and data analysis. It would liberate researchers by giving them the ability to build digital archives of relevant collections. Digital Humanities is interdisciplinary, and while it would be difficult to build a tool that meets every researcher's needs all of the time, a unified platform that allows researchers to efficiently search, compile and analyse materials would present many opportunities to researchers. What elements would such a system need to incorporate to support research and teaching in Digital Humanities?

2.4.1. Mining, Analysing and Exploring Text

A digital tool's ability to mine, analyse and explore text is a valuable element to researchers for the purpose of extracting entities and attributes, and identifying relationships in text documents, thus giving structure to unstructured texts. Mergel et al. (2015) define Text Mining as the process of discovering new information, using algorithms to extract the information from text documents. Janicke et al. (2017) describe close reading and distant reading as two methods with which the visual analysis of a text is done in Digital Humanities. Close reading is the traditional way of reading, which involves reading for meaning and comprehension, without dissolving the structure of the text. Distant reading dissolves the text structure as it applies computational methods to visualize the global features of a text. Close reading techniques visualize text using colour, font size, glyphs and through making

connections to illustrate the structure among textual entities. Distant reading techniques visualize text using heat maps to highlight text snippets, tag clouds to visualize keyword frequency, and maps and timelines to analyse named entities. Slingerland et al. (2017) argue that the strength of distant reading over close reading is the ability for distant reading to pick up trends and patterns in data that may not be visible to human “close” readers. Janicke et al. (2017) list pre-processing approaches in transforming a given text into visualization input formats including tokenization, normalization, sequence alignments, part-of-speech tagging, named entity recognition and, topic modelling.

Text mining and analysis can be useful to researchers wishing to derive quality information from text. As shown in Figure 4, which is a snippet of a Web-based reading and analysis tool, the availability of infrastructure that allows researchers to derive insights from text, whether it be through showing the most frequent terms and trends within a document or a textual overview showing each occurrence of keywords, can help researchers draw significance insights from text documentation.

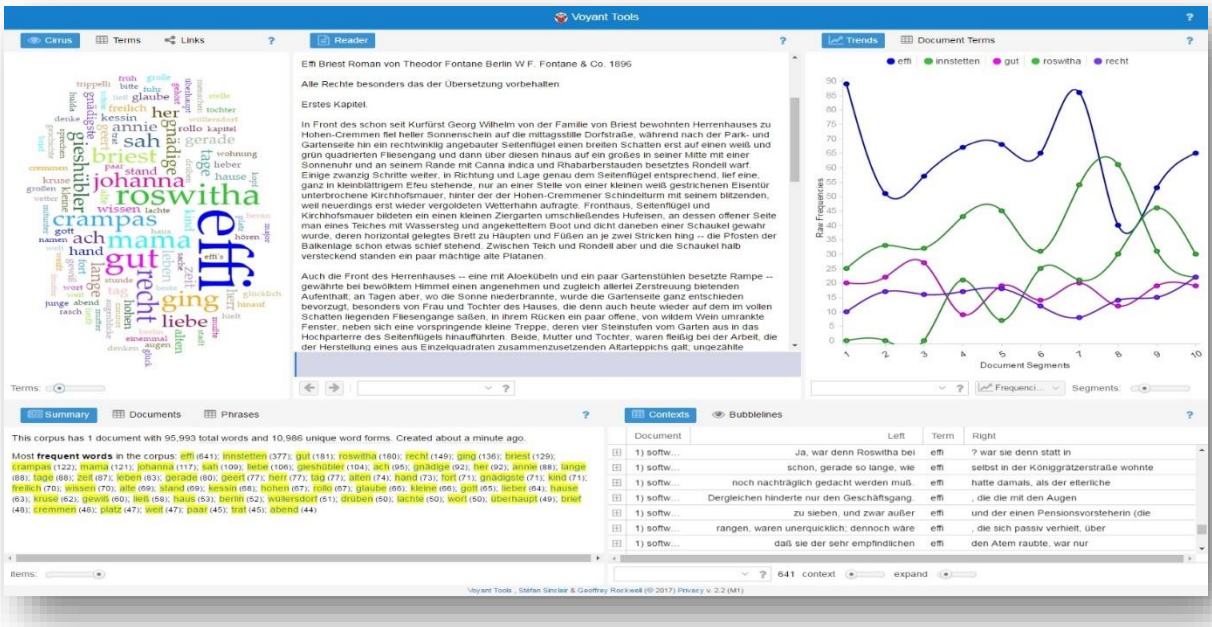

Figure 4: Example of a textual analysis program being used to study a novel (Thomas: 2017)

Tables 1 provides a sample of the tools accessible to researchers for large-scale text mining, analysis and exploration.

Table 1: Tools accessible to researchers for exploring text and literature

Google Ngram Viewer	An online search engine that charts how often a word is used and compares that to other words	https://books.google.com/ngrams
Bookworm for Open Library	Similar to 'Google Ngram Viewer', but searches books in Open Library	
Wordle	Makes word clouds from generated words	http://www.wordle.net/
Voyant Tools	An open-source web-based text reading and analysis tool. It works with various formats including plain text, HTML, XML, PDF, RTF, and MS Word	https://voyant-tools.org/
Natural Language Tool Kit'	A bundle of libraries and programmes for English language processing	https://www.nltk.org/
MALLET	Java-based software for statistical natural language processing. It is good for topic modelling and document classification	http://mallet.cs.umass.edu/
TAPoR	An extensive database of textual analysis tools	http://tapor.ca/home
R	A programming language and statistical software capable of data analysis and visualization	https://www.r-project.org/

2.4.2. Visualizing Data for the Creation and Discovery of Knowledge

If a researcher is up to speed with automatic text analysis and can process complex data in a fast and effective way, what is next? According to Kalpesh et al. (2012), significant research has been done in the area of visualizations for viewing and querying documents, and in the graphical querying and browsing of results. The study mentions popular digital archives for searching through collections, including Archive-It and Internet Archive, and argues that these provide a good user interface for searching through collections. Mergel et al. (2015) define information visualization as a multidisciplinary field that involves representing abstract data, to communicate, explore and analyse information. A good application of Information Visualization should be unusual, informative, efficient and attractive. Sinclair et al. (2013)

advise that, for Humanities, information visualization should accommodate a mix of evidence and argumentation. The study argues that visualization should provide not only ways to interact with the data but should contribute to new and emergent ways of understanding the material being visualized. Jessop (2008) discusses various types of data that can be visualized within digital archives, including spatial data, quantitative data, text data, timelines and 3D visualizations. (1) Spatial data involves the study of spatial relationships using Geographical Information Systems (GIS) software that uses visualization to combine time, space, quantitative data and qualitative data. This type of visualization is used by archaeologists, anthropologists and in the 3D visualization of cultural heritage sites. (2) Quantitative data is the most established form of data used by quantitative analysts for visualization. This data is used by most generic statistical analysis software and applications such as those used in text analysis. Quantitative data is classified into comparative, compositional, distributional, and relational data. (3) Text data makes use of the visualization methods derived from quantitative methods. (4) Timelines, which offer a lot of value to historians, allow one to explore the development of historical events. (5) 3D visualizations, which have focused on visualizing the built environment, have been of value to historians, archaeologists and those interested in understanding the buildings of the past.

These studies provide insight into how data visualization may aid researchers to turn data into information and communicate insights to others. Data visualization assists researchers in the pursuit of knowledge and guides research for information that can ultimately be published. Figure 5 demonstrates the visualization of data using Tableau, a data visualization tool. Data visualization lets researchers see patterns and trends, find correlations between variables and

helps makes comparisons side by side. Researchers can use data visualization to proactively strategize and make it easier for teams to collaborate.

Figure 5: Example of visualizing data using Tableau (Duke Digital Humanities: 2020)

2.4.3.Geographic Information Systems (GIS) Data and Mapping

Geographic Information Systems (GIS) have the potential to expand Humanities research and benefit humanists in their scholarship. Kong et al. (2017) concluded that the introduction of GIS and spatial information to the Humanities could help scholars better understand historical and social sciences and help get scholars started with their digital projects. Gieseking (2018) believes that the production of maps and data visualizations are an important contribution to Digital Humanities and that the critical analysis of the data visualizations and archives that are a result of Digital Humanities research is even more significant. Kallaher and Gamble (2017) state that the study of geography, including work in Geographic Information Systems, allows researchers using various tools of which GIS is one, to identify patterns and trends in spatial

and location data that might not seem connected at first glance. They discuss two types of data used in GIS: geographic data and attribute data. Attribute data refers to information about spatial data. The data can include natural and abstract features, which allow for the management and analysis of geographic space and the derivation of meaning using maps. The visualization and analysis of geographical and location data make use of colours, symbols, geographical statistics, patterns, themes and trends. Over the years, mapping using geographic information systems has become a critical part of Digital Humanities. Janco (2014) believes that an opportunity exists for mapping and spatial analysis tools that allow users to compile and analyse materials from digital archives.

These studies provide insight into the applicability of GIS to quantitative and qualitative types of research that have to include spatial analysis. The GIS service can play an important role in supporting Humanities researchers in their collaborations and learning. These studies can be adapted to form a basis from which to create a general service model about how GIS services can be best developed to engage with Digital Humanities and social science. Figure 6 demonstrates the strength of GIS in the Digital Humanities by showing how manipulations of geographic representations can allow data to be organized in a way that is captivating and easy to process. Representing data in a visual context can help researchers better understand the significance of that data.

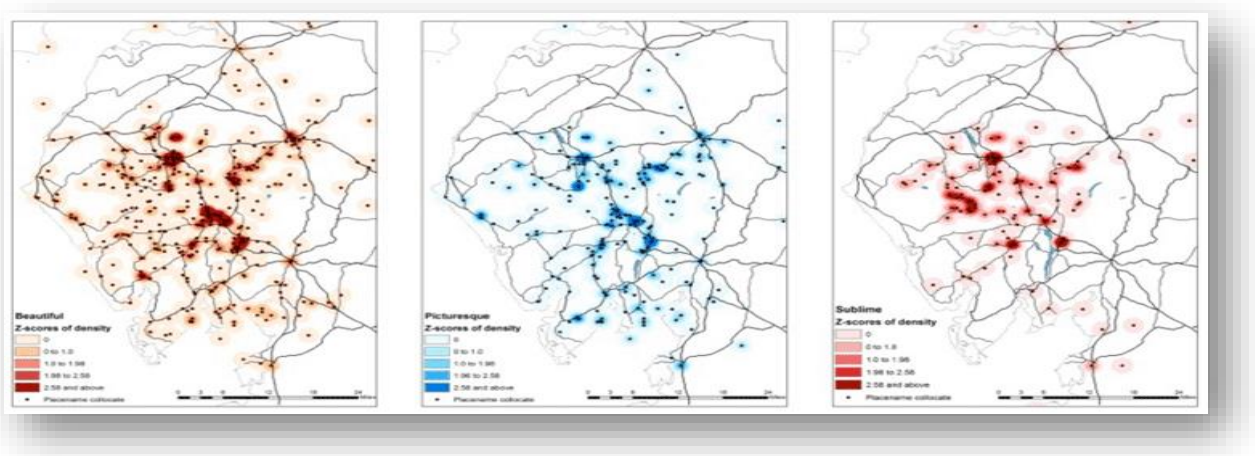

Figure 6: Example of research using text analysis and GIS (Altaweel: 2017)

2.4.4. Building Websites and Publishing Digital Content

It can be challenging for researchers to access traditional non-digital print as it is not always easily sharable, reusable or transformable into different formats. The use of computers to create digital files has allowed for the distribution of scholarship on digital platforms, like websites, making it accessible to more avenues. Velmurugan & Natarajan (2015) define Digital Publishing as a complex process of transferring anything that is in physical print format to a form that can be accessed online or via electronic devices. Digital publishing is traditionally associated with digitized print media such as digital books and peer-reviewed articles, but modern-day digital publishing has broader capabilities, including the ability to publish digital datasets and multimedia. Hunter (2012) says that partaking in digital publishing requires technical infrastructure and expertise in the area of metadata, markup languages, content organization, design and graphics, Web publishing and archiving. By embracing digital publishing, it becomes possible to respond to the changing needs of researchers. According to Howard (2008), building a website to host digital content no longer requires knowledge of the Hypertext Markup Language (HTML), which is the standard markup language for

displaying documents, pictures and multimedia via the Web browser. Languages that make it easier for Web-users to create Website content are replacing HTML. Web-users can by-pass publishers and offer their creations directly to Internet audiences. Kulesz (2011) believes that, despite the infrastructural and human resource-related challenges that exist in the African continent, digital publishing shows potential. The potential of digital publishing will be realized as mobile phone networks expand, the number of Internet users increases and traditional publishing continues to experiment with new technologies.

Digital Humanists favour digital publishing because it allows for information sharing without the constraints imposed by traditional print. However, Digital Humanists must recognize the importance of using digital publishing strategies that fit the context in which they are used. Publishers must adapt their publishing strategies to meet the needs of researchers and the challenges faced by developing countries. Several digital tools are available for digital publishing. Some of these tools are made specifically for Digital Humanities and can be repackaged quite effectively for Humanities research purposes. Table 2 shows several tools accessible to researchers for publishing digital content.

Table 2: Example of tools accessible to researchers for publishing digital content

Scalar	Scalar enables users to assemble media from multiple sources and juxtapose them with their writing in a variety of ways, with the minimal technical expertise required.	https://scalar.me/anvc/scalar/
WordPress	An open-source and free system that allows for creating websites and blogs with good navigation.	https://wordpress.com/
Omeka	A free, open-source web application that allows scholars, museums, and libraries to publish and display heritage objects.	https://omeka.org/
Drupal	A free, open-source system for content management and for managing digital systems	https://www.drupal.org/

2.5. User-Centred Design for Digital Archives

Digital Humanities tools fail to prioritise the needs of the user, resulting in tools and services that are not useful to the user. A study by Gibbs and Owens (2012) suggests that Digital Humanities tools have neglected the typical Humanities user in their design and documentation. It argues that tools tend to be unclear about their limitations and user interface design seems to have been developed as an afterthought for some systems. According to Power et al (2017), bad user experience, unusable functionality, and the inability of a technology to deliver on its goals will render the technology useless. Digital archives developers need to consider user experience as a primary outcome when developing new systems or they put at risk any investment in data and metadata. Unsworth (2003) called for Digital Humanists to do a better job of demonstrating the usefulness of the materials they digitize. According to Borgman (2009), to attract broader interest in Humanities, analytical tools and services have to be more sophisticated, robust, transparent, and easy to use for the motivated Humanities researchers. The tool interfaces must help 'traditional' historians feel more comfortable with new ways of visualizing, analysing, and thinking about sources and data (Gibbs and Owens: 2012).

Fundamental to the user-centred design is the collection of data from users and the incorporation of findings from the data to design tools that users will find useful. Based on Wasson et al. (2016), achieving the goal of user-centred design, which is to optimize the technology around users rather than forcing users to change their behaviour to accommodate the technology, is a collaborative process that involves users, researchers, designers, and other relevant specialists. Agosti & Orio (2012) say user-centred design is critical, even though

user requirements can evolve while users get to know the capabilities of technology. Thoden (2017) states that creating a positive user experience increases the adoption and usage of Digital Humanities tools and services. The study points to ambiguous and inconsistent vocabulary, disregard for graphical conventions, ambiguous system status, missing documentation, missing strategies to avoid mistakes and disregard of the convention for workflows as some of the typical Digital Humanities tools and services design issues.

Efforts to strengthen the usability of a digital tool will ultimately make a difference to the end-user. Tool developers should not disregard the importance of the user interface and documentation. Usability is important because it can mean the difference between the success or failure of a system for managing and preserving digital collections. The goal of a developer should be to make tools that are easy to use and learn. As depicted in Figure 7, User-Centered Design (UCD) involves users in the design process to create highly usable and accessible products.

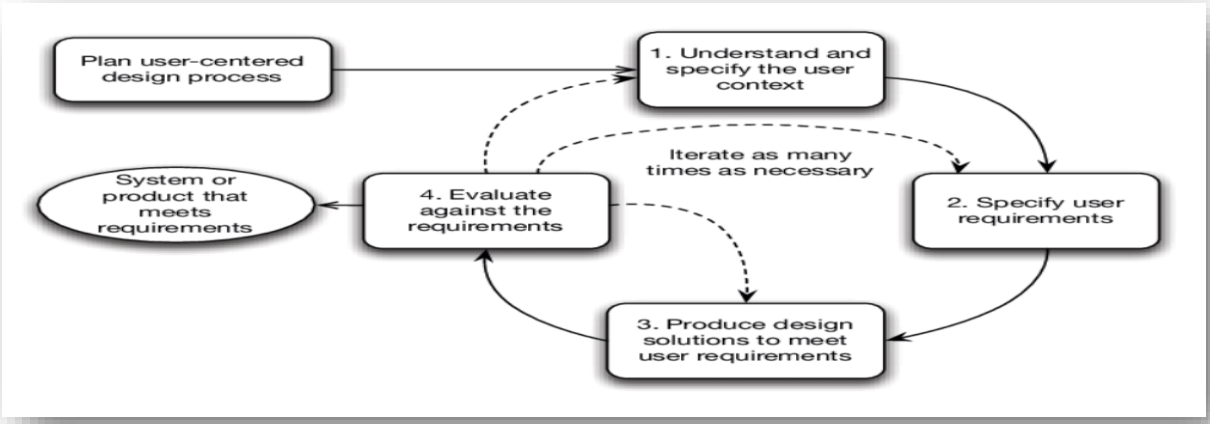

Figure 7: Example of a process of user-centred design (Trenton et al.: 2014)

2.6. Digital Archiving in Poor Countries

The preservation and management of digital objects over the long term are difficult. Many challenges face organisations with the responsibility of managing and preserving digital collections. Asogwa (2012) points to the failure by colonial imperialist to put into practice strong and reliable archival legislation in their colonies as being a contributor to deficient archives and records in African countries. Adu and Ngulube (2017) discuss the risks related to technological obsolescence and point to the rapid pace of technology, the lack of awareness about the significance of digitally documenting historical records, politics-related violence and wars in sub-Saharan Africa as being impeding forces to the development and progress of national heritage institutions in the continent.

Suleman (2019) discusses some challenges for digital archives in developing countries. The challenges in the adoption of Digital Humanities tools in developing countries are partly attributable to limited technical resources, as public institutions typically do not have the technical staff to manage digital repository systems. Limitations in government funding also exacerbate issues, as organizations have to rely on funding from international donors to maintain their archiving operations. The study further points to related issues of Internet access and the priorities of archives. Poor countries for the most part have very limited Internet access and most repositories are Web-based, making access to multimedia files that require much Internet bandwidth difficult. The priorities of small archives are different from those of large repositories as maintaining large complex systems is not a priority for small archives. Phiri and Suleman (2015) argue that the preservation of digital objects is an ongoing process and that routine maintenance and active monitoring is crucial for digital preservation,

but existing funding models cannot be relied upon to ensure this takes place for smaller organizations involved in curation and preservation.

The current design and development approach utilized by developing countries like South Africa for developing digital archives are not producing the types of digital archives suitable for scholarly research and long-term digital preservation. Designing and implementing systems that improve access to digital collections in developing countries requires a complete reinvention of the current design and development approaches and practices and an introduction of innovative strategies that suit the African environment.

2.7. Chapter Summary

The interdisciplinary nature of Digital Humanities and the evolution of Computing has significantly contributed to new approaches to archiving, research, learning, teaching and collaboration, but there continue to be important digitization issues that need to be resolved in computing by Humanists building digital tools for research and learning. The issues include what to digitize, how to represent digitized objects, how to digitize objects accurately, and how to use digital tools to enhance research and learning activities. The adoption of proper digitalization strategies will help refine preservation processes, improve quality and promote consistency in the way records are displayed, so that information becomes accessible at all times, from anywhere, and from any chosen device. While it would be difficult to build tools that meet every researcher's needs all of the time, a unified platform that allows researchers to compile and analyse materials would present many opportunities to researchers. Many challenges face developing countries, making it hard to achieve

sustainable long-term digital preservation. The design and implementation of systems for creating and delivering access to digital collections may require a reinvention of systems, approaches and practices currently used in developing countries and an introduction of innovative strategies that suit the African environment. The next chapter explains the methods used to obtain the desired primary data to develop an understanding of how South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities. It also presents the approach to the desktop study of existing archives and archiving tools in South Africa to determine the extent of features available to users.

Chapter 3: Research Methodology

3.1. Introduction

The purpose of this chapter is to describe the research methodology followed to obtain the desired primary data to develop an understanding of how South African Humanities scholars use digital archives in their research, as well as in teaching and other academic activities. The chapter begins with a feature determination study that provides the sampling frame for the thesis, defines the scope for the survey tool, and is used to uncover trends in digital archives development in South Africa. The feature determination study represents a first step in the development of an exhaustive research questionnaire. This chapter then goes on to provide the rationale and methodological detail for this thesis. It describes the study design, the population and the sample. The chapter also describes the instrument used to collect the data, including methods implemented to maintain the reliability of the research tool. The research survey helps answer this thesis's research question in the context of the technical features of the systems mentioned in the feature determination study.

3.2. Feature Determination

The research began with a desktop study of existing digital archives and archiving tools in South Africa to determine the features available to users. The purpose of the desktop study is to help build a good understanding of the functionalities and services available across Humanities digital archives in South Africa, and to uncover gaps in the functionalities and services offered.

3.2.1. Phase 1: Listing Digital Archives

The first phase in the feature determination study involved constructing a list of all the digital archives available to Humanities researchers in South Africa. The directory used to find archives and repositories was OpenDOAR, a global directory of open access repositories (Sherpa Services: 2020). OpenDOAR hosts different types of archives and repositories, some institutional and other governmental, from countries across the world. Not all the archives found in the directory were ideal for the desktop study. Institutional or governmental repositories that are uniform or based on the same platform or outside of the Humanities fields were not ideal and were therefore left out. This was to keep the list of archives in alignment with the goal of the feature determination study, which is to uncover functionalities across different Humanities digital archives. In addition to the OpenDOAR directory, some South African digital archives were found through references from other archives and repositories found in the OpenDOAR directory. The UCT library and Google Scholar contain high-quality peer-reviewed articles and papers with subject directories containing links selected by subject experts and were additional places to find relevant digital archives and repositories. The archives in Table 3 were included as part of the feature determination study. Four of the archives are not hosted in South Africa, but contain South African content and were included to add to the list of possible features and services.

Table 3: List of archives included in the feature determination study

1	National Archives and Records Service of South Africa	http://www.national.archsrch.gov.za/sm300cv/smws/sm300dl
2	GALA *(Gay and Lesbian Archive)*	https://gala.co.za/archive/
	Anti-Apartheid Movement Archives	https://www.aamarchives.org/
3	Nelson Mandela Foundation - Centre of Memory archive	https://www.nelsonmandela.org/content/page/collections
4	The South African History Archive (SAHA)	http://www.saha.org.za/

5	UWC- Robben Island - Mayibuye Archives	http://mayibuyearchives.org/
6	Wits Historical Papers Research Archive	http://www.historicalpapers.wits.ac.za/
7	Zamani Project Archive (Original site)	https://www.zamaniproject.org/digital-heritage-archive.html
8	Zamani Project Archive (UCT)	https://zivahub.uct.ac.za/zamani_project
9	Desmond & Leah Tutu Legacy Foundation Archive	https://www.tutu.org.za/programmes/archives/
10	UCT Libraries Digital Collections	https://digitalcollections.lib.uct.ac.za/top-level-collection
11	Community Video Education Trust	http://cvet.org.za/
12	Baileys African History Archive	https://www.baha.co.za/
13	O'Malley "The Heart of hope - South Africa's transition from apartheid to democracy")	https://omalley.nelsonmandela.org/omalley/index.php/site/q/03lv00000.htm
14	South Africa: overcoming apartheid, building democracy	https://overcomingapartheid.msu.edu/
15	South African Music Archive Project	https://samap.ukzn.ac.za/
16	SABC Sound Archives	http://web.sabc.co.za/sabc/home/bf/medialibraries/details?id=e1d3bcaa-c1a9-416c-bed8-d7e574112721&title=SABC%20Radio%20Archives
17	SABC Radio Archives	https://sites.google.com/site/sabcmedialib/radio-archives
18	Liliesleaf Farm, Rivonia	https://liliesleaf.co.za/about-us/archive-2/
19	South African History Online Archives	https://www.sahistory.org.za/archives
20	University of South Africa (UNISA) Library, Archives and Special Collection	https://www.unisa.ac.za/sites/corporate/default/Library/Library-services/Archival-and-special-collections/
21	The University of Johannesburg Archives and Special Collections	https://ujcontent.uj.ac.za/vital/access/manager/Index
22	University of Pretoria Archives	https://www.up.ac.za/up-archives
23	University of the Free State Archives for Contemporary Affairs	https://www.ufs.ac.za/supportservices/departments/archive-for-contemporary-affairs-home
24	WITS Private Manuscript and Archival Collections	https://www.wits.ac.za/about-wits/facts-and-figures/central-records-and-archives/collections/
25	International Institute of Social History - Anti-apartheid and southern Africa Collection	https://iisg.amsterdam/en/collections
26	The Freedom Archives	https://search.freedomarchives.org/search.php?view_collection=209
27	African Activist Archive	https://africanactivist.msu.edu/
28	Schaderberg Movie Company	http://www.jurgenschadeberg.com/
29	Columbia University, Oral History Research Office	https://library.columbia.edu/libraries/ccoh.html
30	South African Heritage Resource Agency	https://sahris.sahra.org.za/

3.2.2. Phase 2: Functionalities in DELOS Model

The second phase in the feature determination study involved identifying and listing some of the functionalities commonly supported by digital archives. These include some of the basic functionalities such as searching, sorting and browsing. The literature on Digital Humanities reveals additional and more advanced functionalities that may not be common but available on some archives; these were also included in the list of available functions. Candela et al (2008) list some of the functionalities considered important to users based on the Delos Digital Libraries Reference Model. The Functionality Domain, as shown in Figure 8, captures the processing that can take place on digital resources and the observable activities by users in a Digital Library. In the Delos Digital Libraries model, "Function" is specialised into five areas, namely, Access Resources, Manage Resources, Collaborate, Manage Digital Library and Configure Digital Library resources. Access Resources, Manage Resources & Collaborate functions are directly relevant to users, whilst the Manage Digital library and Configure Digital Library functions are relevant to Digital Library administrators. The Access Resource function includes all activities related to requesting, locating, retrieving and transforming resources. The Manage Resource function includes all activities related to the creation, insertion, deletion, updating, as well as the conversion and transformation of resources. The collaborate function includes all activities that allow users to work together to achieve a common goal.

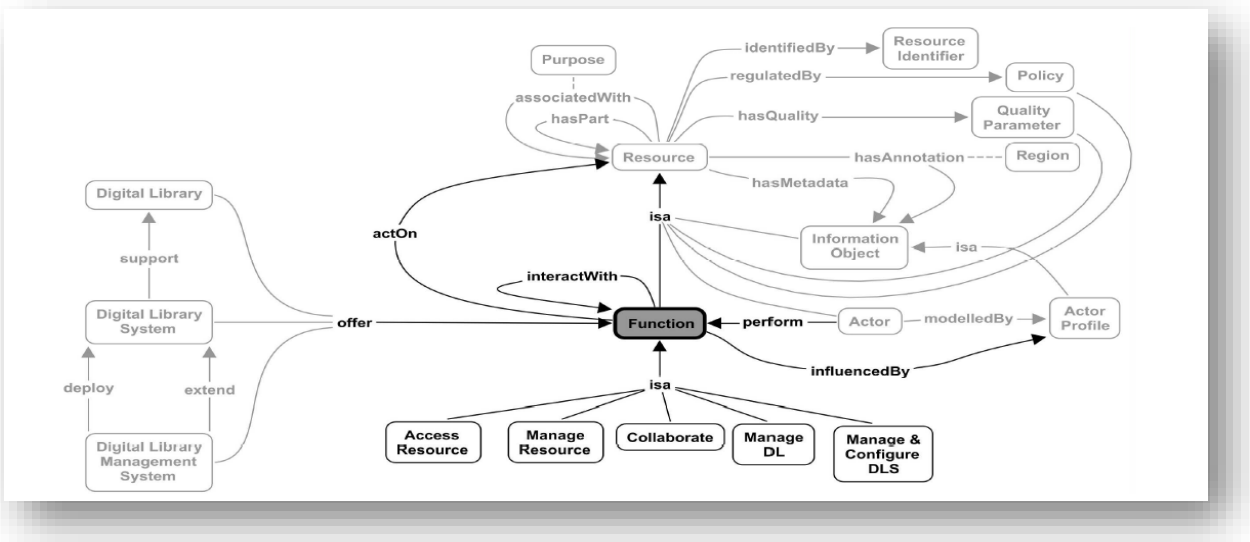

Figure 8: Functionality Domain Concept Map (Candela et al.: 2008)

The following is the categorisation of the functions relevant to users:

Access Resources:
- *Discover*
- *Browse*
- *Search*
- *Acquire*
- *Visualise*

Manage Resource:
- *Update*
- *Validate*
- *Annotate*
- *Disseminate*
- *Compose*
- *Analyse*
- *Transform*
- *Physically Convert*
- *Translate*
- *Extract*

Collaborate:
- *Exchange Information*
- *Converse*
- *Find Collaborator*
- *Author Collaboratively*
- *Social media*

3.2.3. Phase 3: Mapping Process

The mapping process involves visiting the digital archives listed in the first phase of the desktop study and using a matrix structure, such as one seen in Figure 9 and 10, to keep track of which functionalities are available on each archive. Where one of the listed functionalities exists in an archive, this is noted with a tick, and if does not, that is noted using a cross. A matrix structure makes it easy to compare archives across a range of features and to see the

availability of functions across different archives. The matrix structure also shows where different digital archives currently focus their attention on collection building.

		List of Digital Archives		
		Archive 1	Archive 2	Archive 3
Digital Archives common Functionalities	Expected Functionality	X	√	√
	Expected Functionality	√	√	X
	Expected Functionality	X	√	X

Figure 9: Feature determination study matrix (A)

Visiting the different digital archives might reveal functionalities that are not in the list of the common functionalities or those discussed in theory. When such new functionalities are encountered, the matrix is updated accordingly.

		List of Digital Archives		
		Archive 1	Archive 2	Archive 3
Digital Archives common Functionalities	Expected Functionality	X	√	√
	Expected Functionality	√	√	X
	Expected Functionality	X	√	X
	+ New functionality revealed	√	√	√
	+ New functionality revealed	√	√	√

Figure 10: Feature determination study matrix (B)

3.2.4. Phase 4: The Analysis

The analysis phase involves using the matrix structure to analyse the trends uncovered by the feature determination study and to determine which of the functionalities are most prevalent for digital archives in South Africa and which are least common.

3.3. Survey Approach and Design

Ponto (2015) describes survey research as collecting information from a sample of individuals by having them respond to questions. He asserts that this type of research permits for participants' recruitment, data collection and the utilization of instrumentation methods. Questionnaires allow large populations to be assessed with relative ease. The advantage of collecting data through electronic questionnaires is that they allow responses to go directly into a computer database, thus giving way to immediate data compilation, access to larger target populations, visual aids and quick responses. In this exploratory study, primary research data was collected directly from the subjects of the investigation. The primary research method used is an online research survey. The development of the research questionnaire for this study is informed by the feature determination study.

3.4. Research Setting and Data Collection

For this study, an online survey with Humanities researchers in South Africa was used to answer the study's research questions in the context of technical features of systems as demonstrated in the feature determination study. This was a self-administered survey completed via the Internet using LimeSurvey. LimeSurvey is an online open-source survey software for creating and running professional online surveys (Limesurvey GmbH: 2020). It allows respondents to select their responses or provide feedback and makes the collection and analysis of data simple. The system is available to UCT researchers, the data collected remains in the system, and the responses are password protected and are only visible to the researcher. Appendix A lists the survey questions and the response options available to respondents.

3.5. The Study Population and Sample

Banerjee and Chaudhury (2010) define a Population as a set of people with a specialized set of characteristics and a Sample as a subset of the population. The defining criteria for a population may be geographic, demographic, and time-related. The study sample is selected from the population. A study by Kazerooni (2001) points out that the target population consists of all the individuals in the world, with the same characteristics as the sample to which the conclusion of the study is applied. The target population for this thesis are individuals in the Humanities space - academics, postdoc researchers, postgraduates, undergraduates, public and private sector professionals - who make use of digital archives. Given that it is not realistic to perform research on all individuals in the world, a non-random convenience sample is used in this study. A link to the survey was circulated to group mailing lists and to an alumni portal consisting of a spectrum of individuals in the Humanities who make use of digital archives – people of different age groups, different levels of education and different levels of experience using digital resources and digital archives. The goal was to get as many respondents as possible. A link to the survey was circulated via the UCT DSA research invitations mailing list, a general UCT link used to invite UCT students and staff to participate in research surveys. The link was also posted on the Mandela Rhodes Connect portal, a private platform that allows the Mandela Rhodes Alumni community to reconnect via a private platform that only Mandela Rhodes Alumni have access to. The Mandela Rhodes alumni society has many current scholars and alumni in the Humanities. A cold email outreach mailing list was created consisting of Humanities researchers from the University of Pretoria, Rhodes University, Nelson Mandela University, University of the Witwatersrand and the University of KwaZulu-Natal. These universities were selected because they are research

universities and have Humanities researchers' email addresses publicly available. All members of the mailing lists and the alumni portal were invited to respond to the online survey. Respondents from the mailing lists were self-selected, and given that the survey maintained the anonymity of the responses, there was no way to differentiate the characteristics of respondents and non-respondents.

3.6. Reliability of the Research Instrument

The reliability of the research instruments in this study was ensured by eliminating any possible biases in the survey questionnaire by keeping the questions clear and short; leading questions were avoided and definitions for concepts that may be difficult for respondents to understand were provided. The survey questionnaire was kept at a reasonable length and it was ensured that the survey questionnaire measures what it is intended to measure. It was important to test the survey questionnaire for this study before using it to collect data. Pretesting and piloting helped identify questions that do not make sense to participants, or problems with the questionnaire that might lead to biased answers.

The survey questionnaire for this study was pretested with five Humanities researchers representing the target group of this research. The testers were asked to complete the survey one at a time without them seeing each other complete it. The testers completed the survey online and in the same way that it was completed by the respondents. The feedback from the testers was observed to identify where there were hesitations or where they made mistakes. Despite testers not making mistakes, valuable feedback was obtained through the pretesting process, for example, testers indicated that the relationship between some draft survey questions would potentially introduce biases to the study. Testers also suggested that it be

clear when questions speak specifically to South African digital archives and when they do not. A common concern from all testers was that the ranking functionality of LimeSurvey did not always work well on mobile phones. The feedback from the testers was used to improve the survey questions and the layout of the survey questionnaire. The improvements that were made included ensuring that response options to related questions were correlated and that it was much clearer from the outset that the focus of some of the survey questions was on South African digital archives.

3.7. Ethical Considerations

The UCT Faculty of Science Research Ethics Committee approved this research and ethical considerations were considered to ensure that the research was conducted appropriately and following the requirements set out by the Ethics Committee. To comply with ethical considerations in conducting research all survey participants had to consent to participate in the research survey. The Informed Consent Form appears in Appendix B and was included at the beginning of the online research survey. As per the Consent Form, participation in this study was voluntary and participants could exit the survey at any time. The purpose of the survey was explained to the participants. At the end of the survey, respondents were asked to submit their contact information if they wished to receive a copy of the thesis. This required participants to submit personal details and thus reveal their identity. For those who chose to provide personal details, their privacy was protected, and an adequate level of confidentiality was ensured. The study did not make any exaggerations about the aims and objectives of the research and deception in the research was avoided at all costs. All communications concerning the research were done with honesty and transparency. Misleading information, as well as the representation of primary data findings in a biased way, has been avoided.

3.8. Data Analysis

Following the data collection, the data was organised into a format suitable for analysis and then analysed. Close-ended questions were analysed using MS Excel and STATA 15. The questions in the survey are substantially independent, therefore unanswered questions were ignored and each question was analysed independently of other questions. Descriptive statistics were used to analyse the data and the data was visualized using pie diagrams and bar graphs. Content analysis was used to analyse open-ended questions and the emerging characteristics and concepts coming out of the data were visualized using text dashboards.

3.9. Chapter Summary

This chapter described how the research was conducted, including the methodology used to collect data, how the survey respondents were selected, and how the feature determination study informed the development of the research questionnaire. An online survey was used to obtain the primary data for this study. A feature determination study provided the sampling frame and defined the scope of the survey tool. The survey population was chosen from different users of digital archives in Humanities fields and a non-random convenience sample was used. The UCT Faculty of Science Research Ethics Committee approved the research and consent was obtained from the study subjects themselves. Anonymity was maintained throughout the administration of the online questionnaires and report writing. Pretesting ensured that the research tool was accurate, reliable and fit for purpose. The next chapter details the analysis of the survey results and describes the findings of the research.

Chapter 4: Results

4.1. Introduction

This chapter provides analysis and discusses the results of the survey questionnaire, initially discussing the results of the feature determination study. The feature determination study results are followed by an analysis of the aspects of the survey questionnaire under the headings of (i) Sample, (ii) Access Resources, (iii) Manage Resources, (iv) Collaborate and (v) Additional Functionality. The analysis will be followed by a critical discussion of the findings and the links between existing literature and research to determine whether the new data obtained through the survey supports or contradicts existing information. As noted in the previous chapter, the respondents to the online survey are from a spectrum of Humanities fields and given that a convenience sampling method was used to draw respondents from the population, the data does not represent the total population, but only the respondents of the research survey. 102 participants responded to the survey. The data from the survey responses were analysed using a combination of STATA 15 and MS Excel to provide descriptive statistics and to visualize the data in a way that explains the variables examined in this study. Given that the questions in the survey are substantially independent, responses that were left empty by respondents were ignored and each question was analysed independently of other questions. Descriptive statistics used to analyse the results include frequencies, percentages for categorical data and averages.

4.2. Feature Determination Study Results

The feature determination study results show which of the functionalities are most prevalent for digital archives in South Africa and which are least common. Table 4 below shows the results of the study. Where functionality is either available or not available for more than 50% of the archives studied, this is highlighted using the colours green and red respectively.

Table 4: Feature Determination Study analysis

Functionality available for more than 50% of the archives in the study
Functionality not available for more than 50% of the archives in the study

Category	Functionality	% of archives in the study with the functionality	% of archives in the study where function availability is unclear	% of archives in the study without the functionality
Access Resource				
Discover	Allows users to browse collections	81%	0%	19%
Discover	Allows users to search collections	78%	0%	19%
Browse	Collections are organized into a hierarchy	75%	0%	25%
Browse	Browse by author	44%	0%	56%
Browse	Browse by title	66%	0%	34%
Browse	Browse by issue date	44%	0%	56%
Browse	Browse by subject term	47%	0%	53%
Browse	Browse by collection	78%	0%	22%
Browse	Browse by customized fields	41%	0%	59%
Browse	Browse by more than 1 field	53%	0%	47%
Browse	Browse by spatial location	3%	13%	84%
Search	Search by keyword	75%	0%	25%
Search	Search by author	31%	0%	69%
Search	Search by title	44%	0%	56%
Search	Search by issue date	38%	0%	63%
Search	Boolean search	38%	0%	63%
Search	Narrow down search results by date	53%	0%	47%
Search	Search for identical copies of collections/objects	9%	31%	59%
Search	Search by digital content characteristics rather than by keywords (e.g. update frequency, genre, media type)	69%	0%	31%
Search	Safe search to filter adult contents	0%	0%	100%

Search	The distinction between simple searching and advanced searching	34%	0%	66%
Search	Allows users to search bibliographic entries of reference works.	3%	31%	66%
Search	Search results are ranked based on relevancy to the query	72%	0%	28%
Search	Search results are grouped by domain	59%	0%	41%
Search	Each returned result has a content summary	69%	0%	31%
Search	Each collection has a persistent identifier to refer to	100%	0%	0%
Search	Allows users to print out relevant pages	66%	0%	34%
Acquire	Content import from users is supported	19%	16%	66%
Acquire	Generation of collections across several content providers supported	25%	16%	59%
Visualise	Users can move and perform an interaction with collections in 3D space	6%	3%	91%
Visualise	Supports graphical representation of information and data.	0%	0%	100%
Visualise	Supports displaying data related to positions on the Earth's surface (Geographic Information Systems).	3%	0%	97%
Manage Resource				
Update	Actively notifies users about updates to the digital archive. For example, it notifies users every time collections have been updated or changed.	28%	6%	53%
Validate	Validates, for the user's benefit, the quality status of the collections.	0%	0%	84%
Annotate	Allows for annotations/ comments	3%	0%	84%
Disseminate	Alerts users to the import of new Information Objects or Collections	28%	0%	59%
Compose	Allows the user to create new Information objects by reusing existing objects, either in part or as a whole. For example, the user may compose a multimedia album by putting together audio files, song lists and singer biographies.	9%	13%	78%

Analyse	Provides linguistic analysis services to allow users to examine digital objects.	0%	0%	100%
Analyse	Provides qualitative analysis services to allow users to examine digital objects.	0%	0%	100%
Analyse	Provides statistical analysis services to allow users to examine digital objects.	0%	0%	100%
Analyse	Provides scientific analysis services to allow users to examine digital objects.	0%	0%	100%
Transform	Allows different representations of an Information Object (or a set of information objects) and enables the user to perceive information at different levels of abstraction, as desired.	6%	0%	94%
Physically Convert	Provides Text-to-speech conversion services	0%	0%	100%
Physically Convert	Provides speech-to-text conversion services	0%	0%	100%
Physically Convert	Provides texts to spreadsheet or database format conversion services	0%	0%	100%
Physically Convert	Provides data into graphs conversion services	0%	0%	100%
Physically Convert	Provides 3D to 2D conversion services	0%	0%	100%
Physically Convert	Images into colour histograms	0%	0%	100%
Translate	Multilingual information access	0%	0%	100%
Translate	Provides translation services	3%	0%	97%
Extract	Collections can be published on removable media such as CD-ROM/DVD/USB flash drive	47%	0%	38%
Collaborate				
Exchange Information	Allows sharing and exchanging of physical information with peers.	3%	0%	97%
Converse	Allows users to talk to peers and exchange views and opinions through chat services	3%	0%	97%
Find Collaborator	Allows users to locate other *users* of the system that will be eligible for collaboration.	0%	0%	100%
Author Collaboratively	Allows users to collaborate in authoring an Information Object to create a new version of it or collaborative authoring with other users.	0%	0%	100%
Links to social media	Social media links exist	66%	0%	34%

The feature determination study reveals that many of the South African digital archives are missing some of the features for collaboration, accessing resources and managing resources typically considered important to users based on the Delos Digital Libraries Reference Model. Concerning the management of resources, the browse and search functionalities fare better in South Africa than other functions in terms of availability as they are the most common functions in the digital archives included in the feature determination study.

4.3. User Study Sample

The results for "Sample" show the extent to which respondents identify as Digital Humanists, their familiarity and comfort with digital archives, the frequency with which they use digital archives and tools, and the opinion they have of their technical proficiency. The results reveal the respondent's familiarity with methods, practices and objects of study used in Digital Humanities.

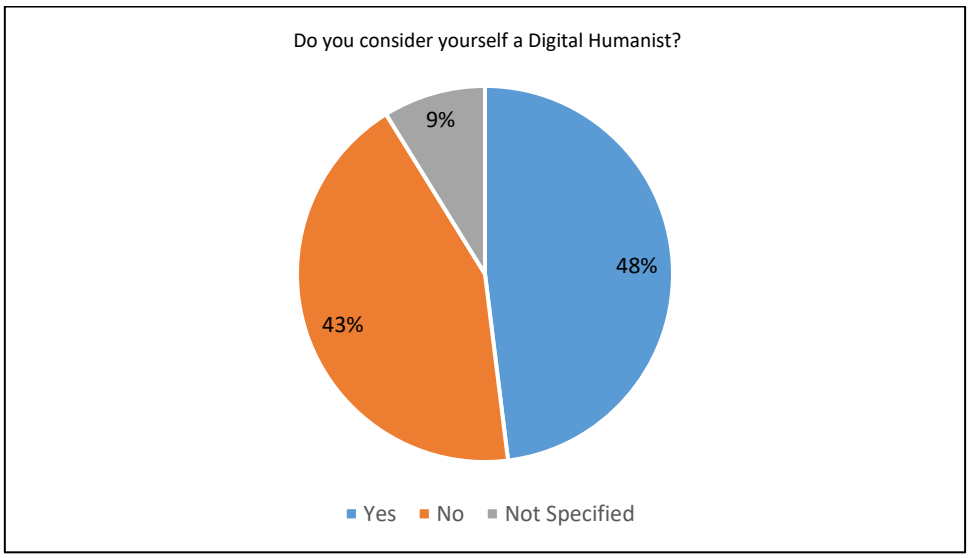

Figure 11: Self Identification as Digital Humanist

As depicted by Figure 11, just less than half (48%) of the respondents consider themselves "Digital Humanist", while 43% do not and 9% chose not to respond to the question. This translates to just less than half (48%) of the respondents perceiving themselves as users of digital resources, complex digital tools and information technologies.

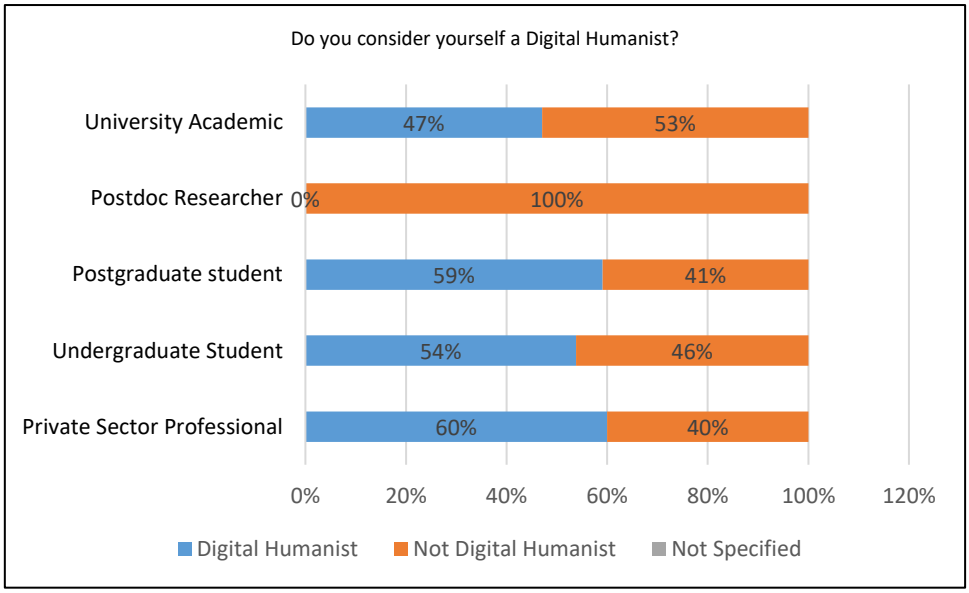

Figure 12: Digital Humanism by occupation

As seen in Figure 12, 100% of those who identified their occupation as "Postdoc researcher" did not identify themselves as "Digital Humanist". 47% of those who identified their occupation as "University academic" identified themselves as "Digital Humanist". Amongst "Postgraduate students", 59% described themselves as "Digital Humanist", compared to 53% "Undergraduate students" and 60% "Private sector professionals". This indicates greater confidence, or greater perceived abilities, among undergraduates, postgraduates and private sector professionals in applying computational methods. This could also be an indicator that there is insufficient confidence among postdoc researchers and university academics in applying computational methods or they underestimate their ability to use digital resources, complex digital tools and information technologies in their scholarship.

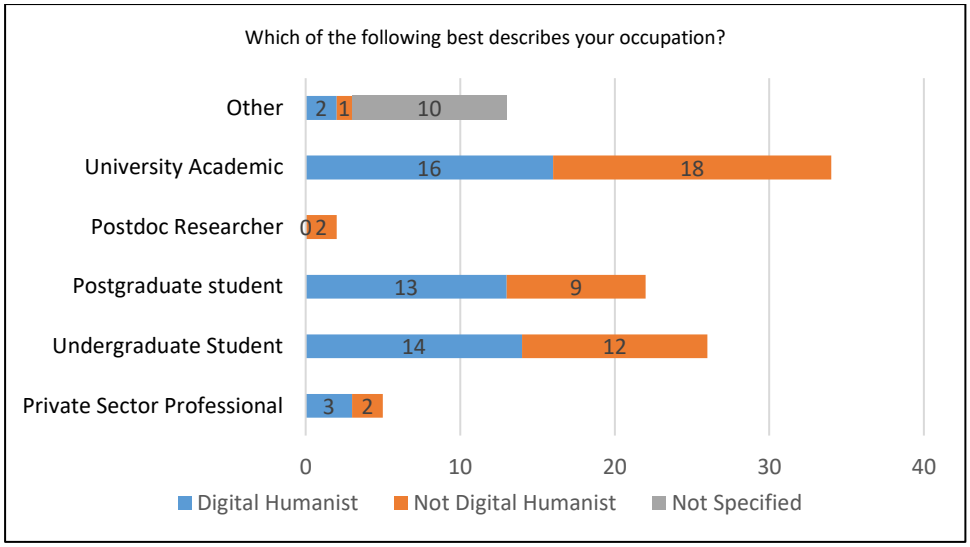

Figure 13: Respondents' occupations

Figure 13 shows that 84 (82.4%) respondents are in academia, including students at undergraduate and postgraduate level. 5 (4.9%) respondents are private sector professionals and 13 (12.7%) described themselves as "Other".

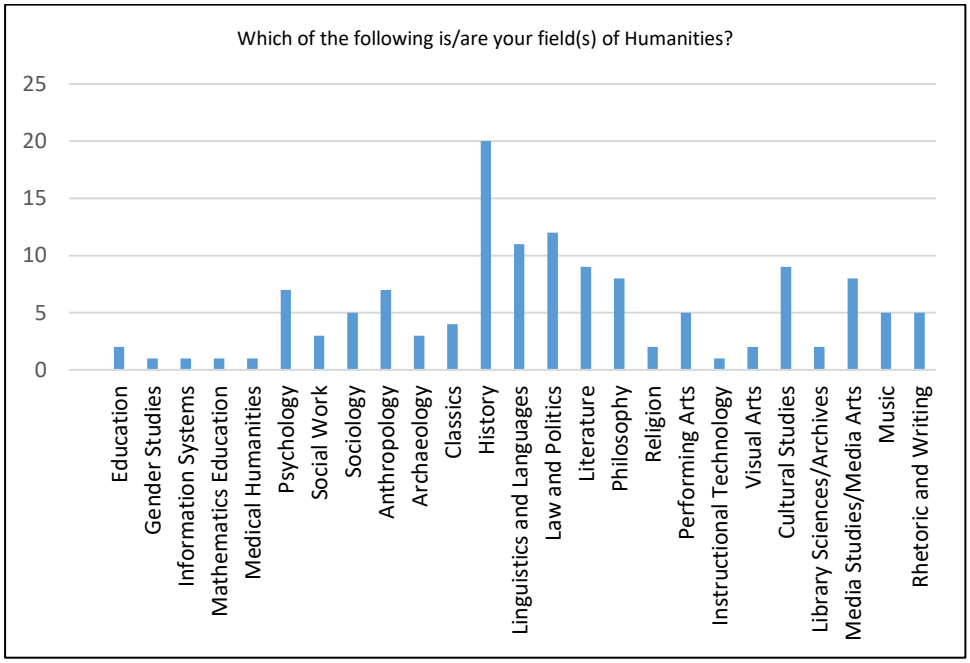

Figure 14: Fields of Humanities

Based on Figure 14, History, "Linguistics and Languages "and "Law and Politics" were the top 3 fields of Humanities mentioned in the responses. Psychology, Anthropology, Literature,

Philosophy, Cultural Studies and Media Studies were other fields that had considerable mentions in the responses.

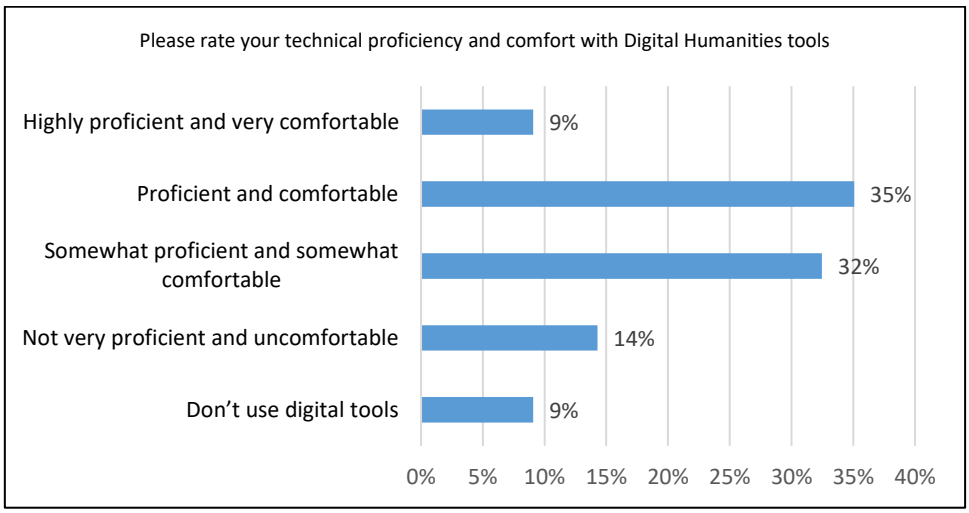

Figure 15: Proficiency and comfort using Digital Humanities tools

As seen in Figure 15, at least 44% of the respondents indicated that they were either highly proficient or proficient and comfortable in using Digital Humanities tools. Meanwhile, 23% were either not very proficient or did not use digital tools at all. Another 32% were somewhat proficient and comfortable. This indicates that 44% of the respondents have acquired the knowledge, experience and skills necessary for using Digital Humanities tools and services, whilst 22% are not confident and require skills training or additional support for the tools they want to use. The 32% that is somewhat proficient and comfortable with using Digital Humanities tools is on route to acquiring the knowledge, experience and skills necessary for using Digital Humanities tools and services.

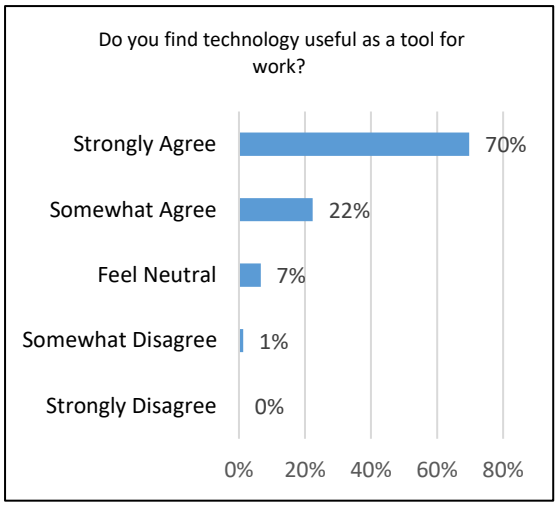

Figure 16: Usefulness of technology

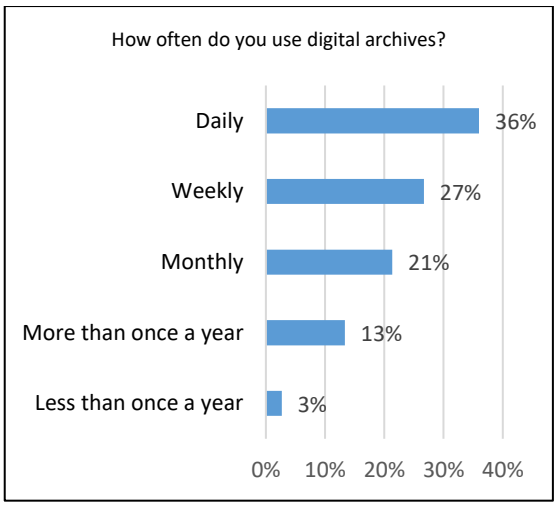

Figure 17: Frequency of digital archives usage

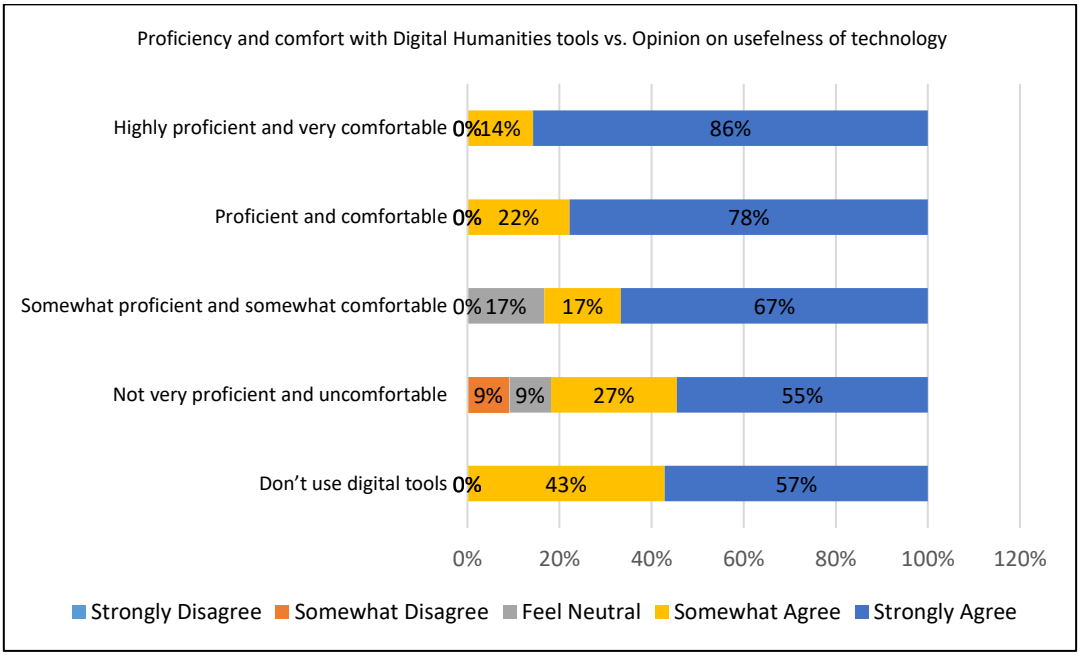

Figure 18: Proficiency using Digital Humanities tools versus the usefulness of technology

As seen in Figure 16, more than 70% of the respondents strongly agreed that they found technology to be a useful tool for their work, indicating a high acceptance of information technology tools among the respondents. Based on Figure 18, of those who indicated that they were "Highly proficient and very comfortable" and those who indicated that they were "Proficient and comfortable", 86% and 78% respectively indicated that they found technology as a useful tool for their work. This might suggest a relationship between "proficiency and

comfort" with "perceptions of usefulness of technology". As shown in Figure 17, 84% of the respondents use digital archives at least once a month, with 36% using them daily.

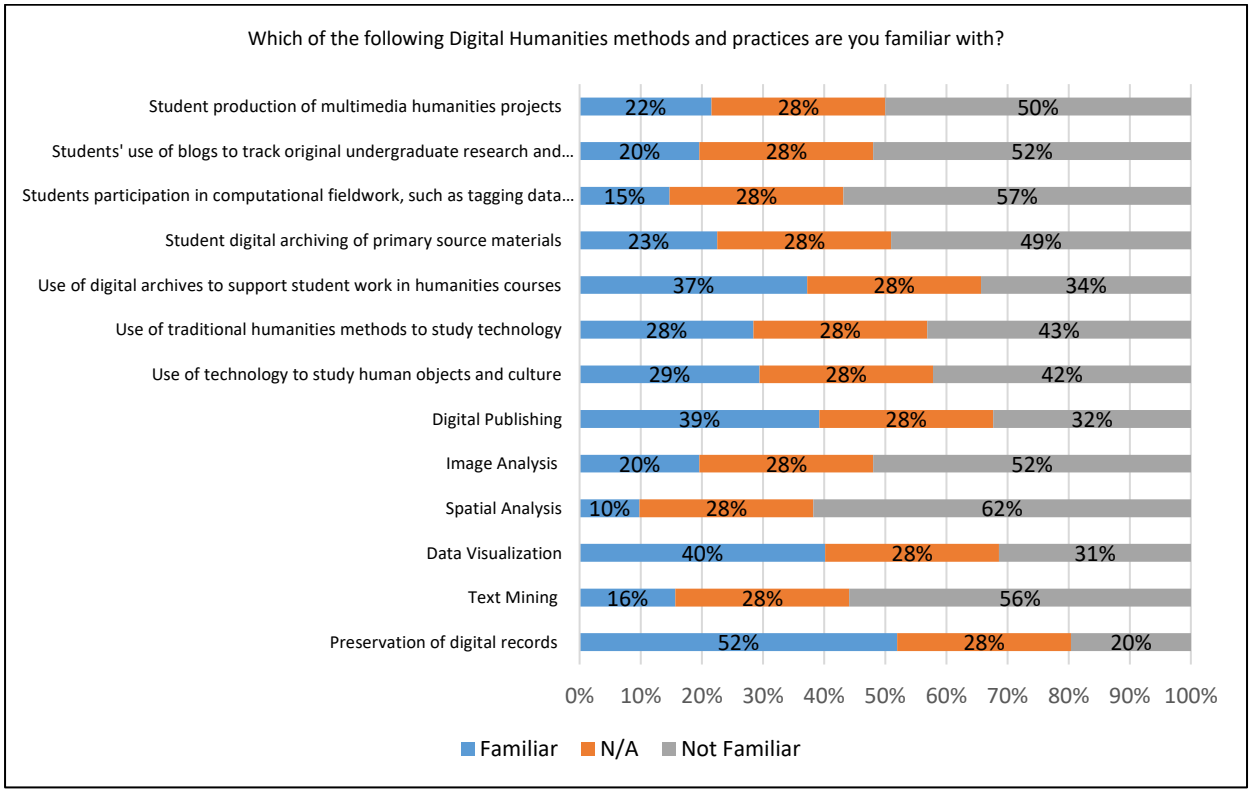

Figure 19: Familiarity with Digital Humanities methods and practices

As shown in Figure 19, on average, 28% of respondents indicated "n/a" when asked about their familiarity with methods, practices of teaching and objects of study used in Digital Humanities, indicating an unwillingness to reveal their familiarity with Digital Humanities methods and practices. On average, 45% of the respondents are "not familiar" with practices of teaching and objects of study used in Digital Humanities. At least 62% of respondents indicated that they are "not familiar" with "Spatial Analysis", followed by 57%, for "Student participation in computational fieldwork, such as tagging data and creating metadata from primary source materials". When adding the 28% average for "n/a" and the 45% average for "not familiar", it is probable that 73% of the respondents are not familiar with practices of teaching and objects of study used in Digital Humanities. On average, 27% of the respondents

are familiar with practices of teaching and objects of study used in Digital Humanities. At least 52% of the respondents indicated that they are familiar with “Preservation of digital records”, 40% of the respondents indicated that they are familiar with “Data Visualization”, and at least 39% of the respondents indicated that they are familiar with “Digital Publishing. These three Digital Humanities methods and practices had the most respondents indicating that they are familiar.

4.4. Access Resources

The results for “Access Resources” are related to digital archives’ ability to provide users with mechanisms for discovering and accessing digital resources. In this section of the survey, respondents were asked about their satisfaction with the South African digital archives they use regularly. They were asked about their satisfaction with the archives’ ability to make content easy to find and accessible to users. They were also asked to rank some of the important functionalities found in digital archives.

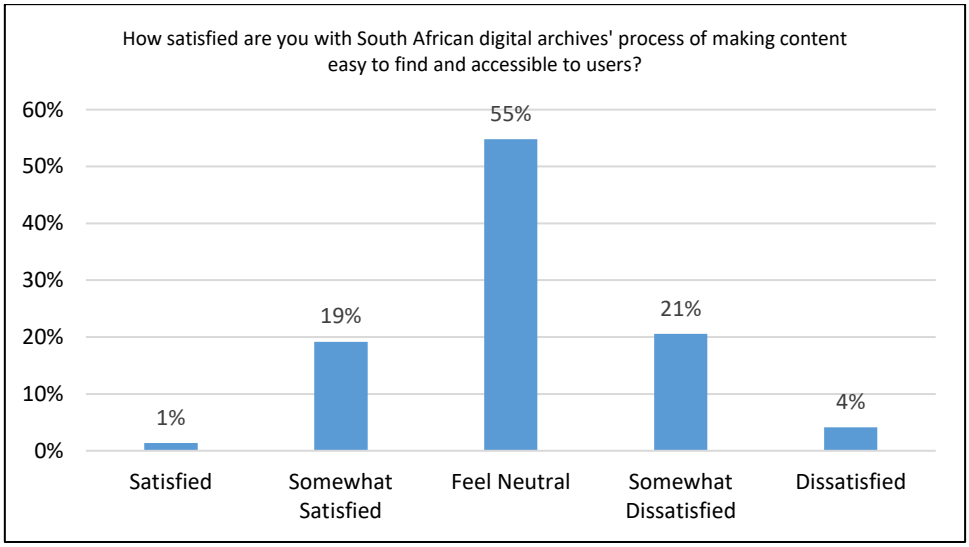

Figure 20: Satisfaction with content accessibility

As seen in Figure 20, only 20% of respondents are satisfied with South African digital archives' process of making content easy to find and accessible to users, with the remainder of respondents neutral or dissatisfied. This indicates a need to improve the services for making resources discoverable and accessible to researchers and an opportunity for the optimisation of the digital archives.

Table 5: Functionalities ranking

Please rank the following functionalities found in digital archives in order of importance from 1 to 5	Position 1 (Highest)	Position 2	Position 3	Position 4	Position 5 (Lowest)
Search	63%	26%	7%	0%	4%
Browse	31%	41%	22%	6%	0%
Acquire	8%	27%	35%	17%	13%
Create	5%	5%	12%	24%	53%
Visualise	0%	4%	26%	52%	18%

Search and Browse ranked highest in terms of importance to respondents as can be seen in Table 5. Search ranked in the top 2 positions by more than 89% of the respondents, followed by Browse, which ranked in the top 2 by 72% of respondents. Only 4% and 6% of respondents ranked Search and Browse respectively in the bottom two positions. Create and Visualise ranked the lowest in terms of importance to respondents. Create ranked in the bottom 2 positions by 77% of the respondents. Visualise ranked in the bottom 2 positions by 70% of respondents.

4.5. Manage Resources

The results for "Manage Resources" are related to functions that support the production, withdrawal or update of digital resources. In this section of the survey, respondents were asked to evaluate several statements related to functions for managing resources and indicate whether they would like to have such functions available in digital archives.

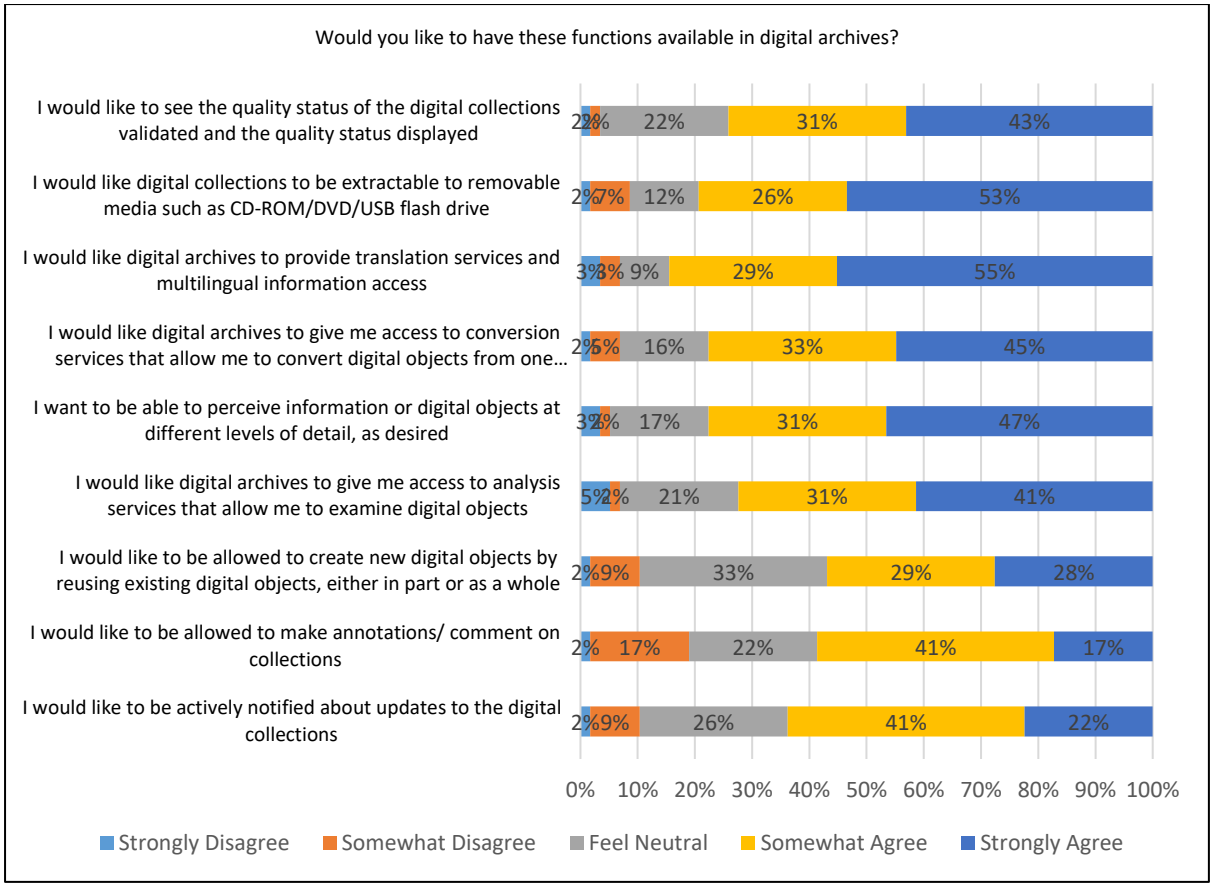

Figure 21: Preferences for digital archives resources

At least 57% of respondents agree or strongly agree with each of the statements in Figure 21. Translation Services are the highest scoring, with more than 84% of the respondents agreeing or strongly agreeing that they would like archives to provide translation services and multilingual access. Extraction of digital collections to removable media such as CD-ROM/DVD/USB flash drive was the preference respondents scored the second highest. Annotations/comments on collections had the lowest scoring with just over 59% of respondents agreeing or strongly agreeing that they would like to be allowed to make annotations and comments on archival collections. Similarly to annotations/comments on collections, respondents scored active notifications about updates to collections low compared to other preferences.

4.6. Collaborate

The results for “Collaborate” are related to the functions that support users in information sharing, communication and collaborating effectively and efficiently with peers. In this section of the survey, respondents were asked how often they would use collaboration functions if digital archives offered them.

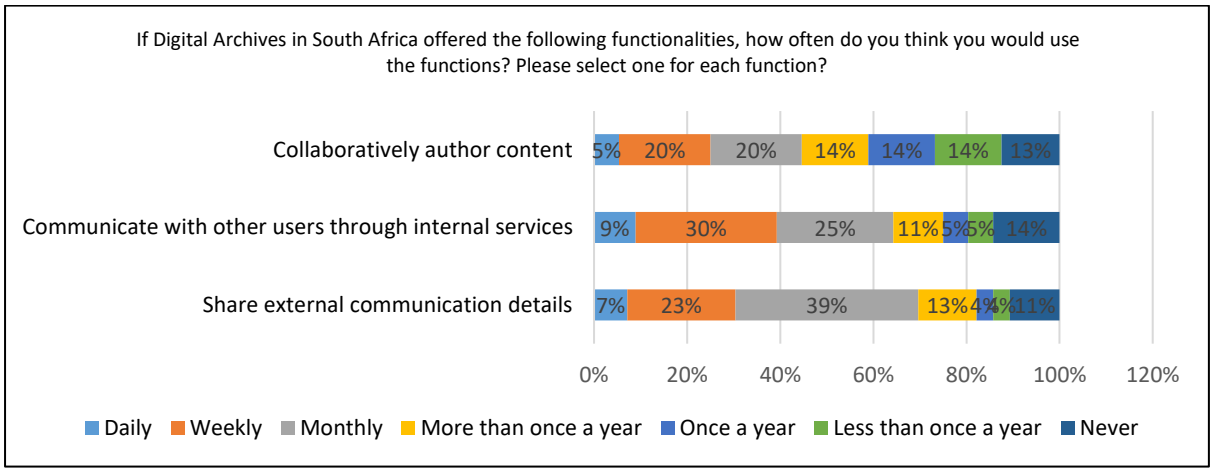

Figure 22: Frequency of collaboration functions usage

As shown in Figure 22, “Sharing external communication details” stands out, with more than half of the respondents indicating that they would use it monthly. “Collaboratively author content” is spread more evenly, indicating that there would not be much variation in how respondents use it throughout the year. Most respondents would communicate with other users through services provided by digital archives weekly.

4.7. Additional Functionality

The results for “Additional Functionality” are related to the overall experience users have had with digital archives, expressed statistically. In this section of the survey, respondents were

asked about their overall experiences with South African digital archives and for their opinions on additional functionality that should be added to digital archives and tools.

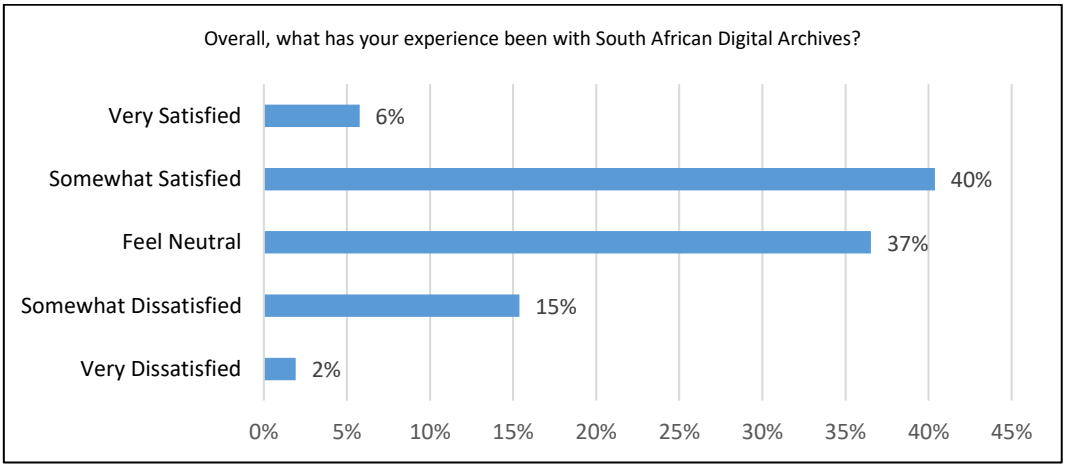

Figure 23: Experience with South African digital archives

Figure 23 shows that at least 54% of respondents are either neutral or dissatisfied with South African digital archives.

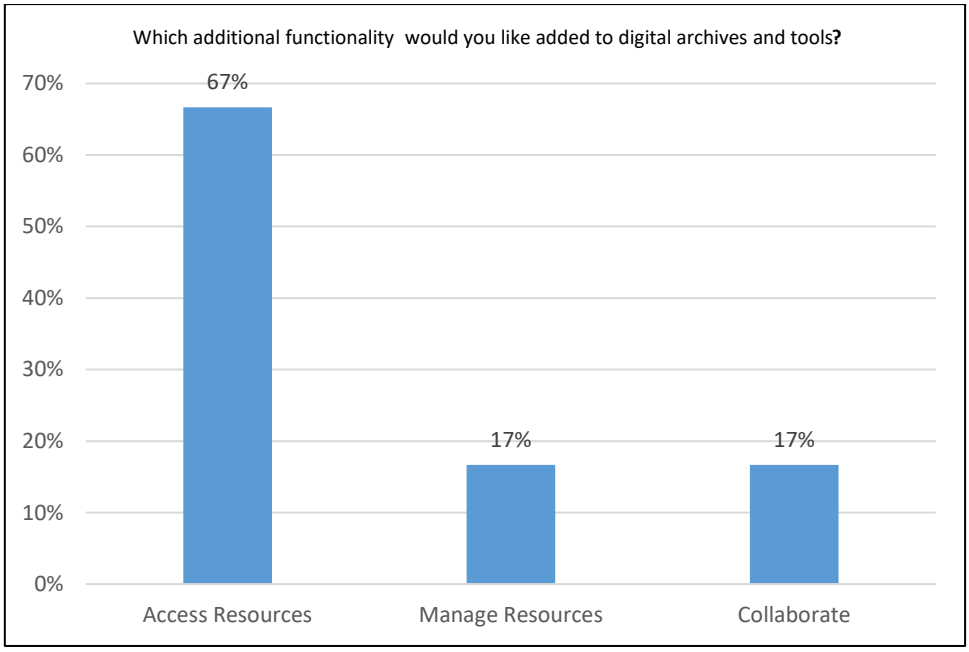

Figure 24: Preference for additional functionality

Besides those who had no comment, the Access Resources functions were the most mentioned by respondents in terms of the need for improvement, as seen in Figure 24. This

supports the findings of rankings, where Search and Browse functionalities were ranked 1 and 2 in terms of importance. Respondents suggested easier navigation, improved search, cross-referencing, linked object catalogue details, optical character recognition, automated suggestions, tutorials, access to search engines from local and known platforms, and metadata correlation across digital archives as additional "Access Resources" functions they would add to archival and non-archival tools. "Manage Resources" & "Collaborate" functions were the least mentioned by respondents in terms of the need for improvement, each only getting 17% of mentions. Respondents suggested analysis and access to a variety of language options as additional "Manage Resources" functions they would add to archival and non-archival tools. Respondents also suggested functionalities that connect the database with researchers involved in similar projects and a chat help function as additional "Collaborate" functions they would add to archival and non-archival tools.

4.8. Chapter Summary

This chapter gave a detailed analysis of the survey results and described the findings of the research. The feature determination study shows that South African digital archives are missing some of the features for collaboration and the management of resources, but the functions for accessing resources such as browse and search are very common. Many of the users of South African digital archives, as represented by the survey respondents, generally do not identify as Digital Humanists, are not familiar with methods, practices of teaching, and objects of study used in Digital Humanities, but use digital archives frequently, find the technology useful in their work, and consider themselves technically proficient in the use of technology. This coincides with Terras' (2016) statistic that in 2014 nearly 50% of faculty members in the Humanities had created or managed digital resources and findings from

Muhammad (2010) that Humanities scholars pay attention to electronic technology, and despite the problems they face in retrieving and using digital resources, find the technology useful in their work. Survey respondents also indicated that they are generally not satisfied with the digital archive's processes for making content easy to find and accessible. This is not surprising considering the challenges for digital archives in developing countries as discussed by Suleman (2019) and the lack of awareness about the significance of digitally documenting historical records as indicated in Adu and Ngulube (2017). Concerning the management of resources, which relate to the production, withdrawal or the update of digital resources, respondents generally want to be able to efficiently extract information from the archives and create resources from existing digital objects but want to do it in languages and formats and levels of abstraction desirable to them. This is an issue fundamental to user-centred design as discussed on Wasson et al. (2016) and Thoden (2017). Creating a positive user experience increases the adoption and usage of Digital Humanities tools and services and part of that includes resolving issues in tools and services design. Respondents are generally not satisfied with South African digital archives' ability to make content easy to find and accessible to users. Improvements in Search and Browse functions, the addition of translation services and multilingual access, and the ability to share external communication details with peers are the areas respondents would like to see improved. This corresponds to Sastry and Reddy (2009) suggestion that digital archives should provide efficient search and browse functionality with user-friendly interfaces, thus bringing about advances in "user-friendly" computer systems. The following chapter briefly revisits the rationale for the study, presents a summary of the main findings and gives recommendations for future research.

Chapter 5: Conclusion

5.1. Introduction

The study aimed to investigate how South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities. The research question was formulated as follows: "How do South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities? The first chapter of this study provided a summary of the research problem and explained the purpose and significance of the research and how the study was carried out. The second chapter was a literature review, which gave a historical development of Digital Humanities and digital archives, identified the research that has already been completed in the area of Digital Humanities and archiving, and provided an analysis of the current information relevant to the research topic. The third chapter explained the methods used to gather information and data to answer the research question and the fourth chapter provided analysis and discussed the results. The chapter aims to present the conclusions drawn from the results of the analysis of the survey questionnaire and the feature determination study, and to make recommendations for further research.

5.2. Addressing the Problem Statement

This thesis addressed the problem statement in part by attempting to compile a list of digital archives available in South Africa and by analysing the digital archives landscape in South Africa to understand the functionalities digital archives have to offer the user. The research methodology part of the thesis engaged Humanities researchers, asking them about their experiences, knowledge areas, preferences, skills and usage of digital archives. This was to

gain an understanding of their general perception of technology and how it is integrated to their daily research activities. The thesis also looks at models and theories developed to guide digital archives and digital library developers towards building tools that will meet the information needs of various users.

5.3. Answering the Key Research Question

There was one key research question in this thesis, which is "How do South African Humanities scholars use digital archives in their research as well as in teaching and other academic activities?" The question has been answered as follows:

South African Humanities researchers use digital archives to access and interact with a wide range of collections. Their activities range from discovering, searching, browsing and gathering content from multiple digital collections to incorporating digital content into tools or social media platforms of their choosing. This is evident in researchers wanting to be able to easily access consistent information that is of high quality and their belief that digital archives have to provide improved usability and better tools for collaboration and for navigating resources. Having digital content as accessible and open as possible, mechanisms for discovering and exposing content to users, capabilities for greater interactivity with digital collections, and collaborations are important to researchers.

Researchers have different preconceived ideas about digital archives, with some still preferring to use traditional written documents and library books. Online archives are

perceived as more complicated by some researchers partly due to the researcher's insufficient confidence in applying computational methods or the lack of ability to use digital resources, complex digital tools and information technologies in scholarship. However, when they do use archives, both traditional and digital, researcher's activities mostly concern supporting student work in Humanities courses and activities that involve the preservation of digital records, data visualization and digital publishing. Despite South Africa Humanities researchers being proficient and comfortable using technology, and despite them finding the technology useful for their work, not many use digital devices to make critical and theoretical observations, as many researchers are not familiar with many of the methods, practices of teaching and objects of study used in Digital Humanities.

Despite South African Humanities researchers wanting to access digital materials in an organized way, they currently do not use digital archives to create different workspaces for individual projects, use the functionality to allow bookmarks and annotations to be categorized and ordered, or copy and move bookmarks and annotations. This is partly due to such functionality not being available on many digital archives. Unfortunately, some researchers do not show a great need to make annotations and comments to collections and have little desire to receive notifications about updates to collections.

South African Humanities researchers use computing technologies to evaluate digital cultural materials and access resources for research, learning and teaching, but the uptake of digital archives is limited by the unavailability of tools, products and services that are easily adapted to incentivise new users to use and invest in them. Despite the current limitations,

researchers want to be able to create new digital objects by reusing existing digital objects, access analysis services to examine digital objects, perceive information at different levels of detail, convert digital objects from one format to another, access translation services and extract content to removable media.

Collaboration is important to South African Humanities researcher's digital scholarship, so scholars do not only compete but are keen to collaborate as well. This is evident because researchers put value in sharing external communication details, communicating via archives' services and collaborating in creating content effectively and efficiently with peers. The frequency with which they use these services varies, with collaboration in creating content being the services they would use consistently throughout the year, given the availability of the service in digital archives. Unfortunately, many of the digital archives available to humanities researchers do not have many of the features for collaboration.

Overall, it is clear that the usage of archives and functionalities of archives vary widely. Users have stronger preferences for some tools, especially tools for basic discovery and for supporting personal and collaborative research, with many users considering the existing support to be inadequate. In terms of advanced functionalities, users are interested in these to varying levels, but the lack of support and the inadequate support for basic functions means that these are still somewhat speculative.

5.4. Contributions of the Study

This thesis contributes to theory, methodology and practice in Digital Humanities. Both the theoretical and empirical findings of the study contribute to the understanding of digital archives development for the Humanities and the preferences researchers have for digital archives features. This study also contributes to the understanding of the question of how researchers in developing countries interact with digital archives and tools. Another contribution is in the experience gained using a feature determination study strategy and the techniques applied for digital archive feature data collection and analysis. The experience might be useful for other studies in Digital Humanities, especially in the context of developing countries where digital archives are still in their infancy. Several models for developing digital archives are discussed, giving developers of digital archives in developing countries a basis from which to work when building new archives and archiving tools. The information gathered from users will aid the development of tools that are potentially more effective, appropriate for the South African environment and widely used by South African researchers. Besides researchers, this topic is of interest to individuals involved in Information Technology related innovations in Digital Humanities and the development of Digital Archives. The study is also of interest to public sector and private sector professionals, managers, and decision-makers responsible for the implementation of Digital Humanities initiatives and the development of Digital Archives. The topic might also be of interest to agencies involved in addressing Digital Humanities issues. The study's focus of attention is therefore of interest to the academic community and it is a developing area of research.

5.5. Limitations of the Study

Given that the Covid-19 pandemic shut down several in-person activities and forced learning and work to take place online, the online method of data collection was the best possible option for this study. Keeping participants fully engaged for an extended period in an online survey is a challenge, so the questions asked in the survey had to be limited to account for the participant's attention span. This means that not all questions that could have been asked were asked. The questionnaire was sent out towards the end of a challenging year, where time was a constraint for many respondents. This contributed to a limited number of completed questionnaires being received. Given that the questionnaire was online, many potential respondents either ignored or did not see the request to complete the questionnaire and repeated requests to complete the survey were perceived as annoying, which also contributed to the low response rate. Given that a convenience sample was used for this study, the data does not represent the total population, but only the respondents of the study.

5.6. Recommendations for Future Research

As an extension to this thesis, a study with a larger sample size drawn from participants actively involved in Humanities research can be conducted. This study would likely produce improved results, especially if the study is carried out over a period that allows for significant data collection. The survey questionnaire can be designed so that questions related to the management of digital archives come across more strongly. A probability sampling method rather than convenience sampling can be used so that the study is more representative of the entire Humanities research community. The feature determination study was carried out

from a limited number of repositories and digital archives available in South Africa. Looking into repositories and digital archives across the African continent would provide an alternative sampling frame and broaden the scope of the survey tool. Future work could include doing similar studies in other disciplines, implementing interventions that include new archives, features, training, and measuring the impact of such interventions.

References

1. Adu, K. K., & Ngulube,K., 2017. Key threats and challenges to the preservation of digital records of public institutions in Ghana, Information, Communication & Society, 20:8, 1127 1145, DOI: 10.1080/1369118X.2016.1218527

2. Agosti, M., and Orio, N., 2012. “User Requirements for Effective Access to Digital Archives of Manuscripts.” Journal of Multimedia 7.2: 217–222.

3. Altaweel, M., 2017. GIS and Digital Humanities, GIS Lounge. Available at: https://www.gislounge.com/gis-digital-Humanities/

4. Asogwa, B.E., 2012. "The challenge of managing electronic records in developing countries: Implications for records managers in sub Saharan Africa", Records Management Journal, Vol. 22 No. 3, pp. 198-211. https://doi.org/10.1108/09565691211283156

5. Banerjee, A., & Chaudhury, S., 2010. Statistics without tears: Populations and samples. Industrial psychiatry journal, 19(1), 60–65. https://doi.org/10.4103/0972-6748.77642

6. Baucom, E., 2019. "A Brief History of Digital Preservation". Mansfield Library Faculty Publications. 31. https://scholarworks.umt.edu/ml_pubs/31

7. Habert, B. C., 2020. Building together digital archives for research in social sciences and Humanities. https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.404.6912

8. Howard, RG., 2008. ‘The Vernacular Web of Participatory Media’, Critical Studies in Media Communication, vol. 25, no. 5, pp. 490–513, https://search-ebscohost-com.ezproxy.uct.ac.za/login.aspx?direct=true&db=ufh&AN=34950039&site=ehost-live

9. Borgman, C., 2009. The digital future is now: A call to action for the humanities. Digital Humanities Quarterly, 3(4). http://digitalhumanities.org/dhq/vol/3/4/000077/000077.html

10. Busa, R., 1980. The Annals Of Humanities Computing: The Index Thomisticus. [ebook] Computers and the Humanities, pp.Vol. 14, No. 2 (Oct., 1980), pp. 83-90. Available at: <https://www.jstor.org/stable/30207304>

11. Byrd, D. and Crawford, T., 2001. Problems Of Music Information Retrieval In The Real World. [ebook] University of Massachusetts Amherst. Available at: <https://scholarworks.umass.edu/cs_faculty_pubs/82>

12. Candela et al., 2007. The DELOS Digital Library Reference Model. Foundations for Digital Libraries (Version 0.96).

13. Day M., 2001. Metadata for Digital Preservation: A Review of Recent Developments. In: Constantopoulos P., Sølvberg I.T. (eds) Research and Advanced Technology for Digital Libraries. ECDL 2001. Lecture Notes in Computer Science, vol 2163. Springer, Berlin, Heidelberg

14. Dougherty, C.K., 2007. Ils.unc.edu. Available at: https://ils.unc.edu/digccurr2007/papers/dougherty_paper_4-2.pdf

15. Dowson, R., 2020. Digital Humanities Research Guide: Tools & Methods | SFU Library. [online] Lib.sfu.ca. Available at: <https://www.lib.sfu.ca/help/publish/dh/dh-guide/tools>

16. Duke Digital Humanities., 2020. Tableau Available at: https://digitalHumanities.duke.edu/tools/tableau

17. DigitalHumanities.berkeley.edu., 2020. Choosing A Platform For Your Project Website | Digital Humanities. [online] Available at: <https://digitalHumanities.berkeley.edu/blog/13/12/04/choosing-platform-your-project-website>

18. Eiteljorg, H., 2001. CSA Newsletter, Fall '01: Computing In The Archaeological Curriculum. [online] Csanet.org. Available at: <http://csanet.org/newsletter/fall01/nlf0104.html>

19. Garaba, F., 2014. Preparing collections for digitisation : the case of religious archives in Pietermaritzburg, KwaZulu-Natal, South Africa : original research. [Online] 35 (1), 1–5.

20. Gibbs F., Owens T., 2012. “Building Better Digital Humanities Tools: Towards Broader Audiences and User-centred Designs”, Digital Humanities Quarterly, vol. 6, no. 2.

21. Gieseking, J. J., 2018. Where are we? the method of mapping with GIS in digital Humanities. American Quarterly, 70(3), 641-648. doi:http://dx.doi.org.ezproxy.uct.ac.za/10.1353/aq.2018.0047

22. Glick, K. . et al., 2006. ‘The ingest and maintenance of electronic records: moving from theory to practice’, in Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL ’06). [Online]. June 2006 IEEE. pp. 359–359.

23. Given, L. and Willson, R., 2018. "Information technology and the Humanities scholar: Documenting digital research practices", Journal of the Association for Information Science and Technology, 69(6), pp. 807-819. doi: 10.1002/asi.24008.

24. Green H. E., Lampron P., 2017. User Engagement with Digital Archives for Research and Teaching: A Case Study of Emblematica Online., Libraries and the academy, Vol. 17, No. 4 (2017), pp. 759–775.

25. Hedstrom, M., 1995. Preserving Digital Information. http://www.ukoln.ac.uk/services/papers/bl/rdr6238/paper.html#hedstrom

26. Huron, D., 1999. The New Empiricism: Systematic Musicology In A Postmodern Age. Ohio State University, p.Ernest Bloch Lectures, Lecture 3: Methodology.

27. Hunter, B., 2012. The Effect Of Digital Publishing On Technical Services In University Libraries. [ebook] The Journal of Academic Librarianship, pp.Volume 39, Issue 1, January 2013, Pages 84-93. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0099133312001449>

28. Ioannidis et al., 2008. DelosDLMS: From the DELOS vision to the implementation of a future digital library management system. International journal on digital libraries. [Online] 9 (2), 101–114.

29. Isah et al., 2013. “Digital Libraries: Analysis of Delos Reference Model and 5S Theory,” Journal of Information Science Theory and Practice. Korea Institute of Science and Technology Information, 1(4), pp. 38–47. doi: 10.1633/JISTAP.2013.1.4.3.

30. Janco, A., 2014. Archival Research Using Digital Archives. [online] Dissertation Reviews. Available at: <http://dissertationreviews.org/archives/8672>

31. Janicke, S., Franzini, G., Cheema, F. and Scheuermann1, G., 2017. Visual Text Analysis In Digital Humanities. John Wiley & Sons Ltd, pp.Volume 36 (2017), number 6 pp. 226–250.

32. Jessop, M., 2008. Digital visualization as a scholarly activity, Literary and Linguistic Computing, Volume 23, Issue 3, Pages 281–293, https://doi.org/10.1093/llc/fqn016

33. Jones, T., Baxter, M. and Khanduja, V., 2013. "A quick guide to survey research", The Annals of The Royal College of Surgeons of England, 95(1), pp. 5-7. doi: 10.1308/003588413x13511609956372.

34. Padia, K., Alnoamany, Y., and Weigle, M., 2012. Visualizing digital collections at archive-it. 15-18. 10.1145/2232817.2232821. Conference: Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries

35. Kallaher, A., and Gamble, A., 2017. GIS and the Humanities: Presenting a path to digital scholarship with the Story Map app, College & Undergraduate Libraries, 24:2-4, 559-573, DOI: 10.1080/10691316.2017.1327386

36. Kazerooni, E. A., 2001. Fundamentals of Clinical Research for Radiologists. American Journal of Roentgenology 2001 177:5, 993-999

37. Klein, J., 2014. Interdisciplining Digital Humanities: Boundary Work in an Emerging Field. doi: 10.3998/dh.12869322.0001.001

38. Klein, S., 2020. Research Guides: Interactive Technology And Pedagogy (ITP): DH Tools. [online] Libguides.gc.cuny.edu. Available at: <https://libguides.gc.cuny.edu/c.php?g=159584&p=1044900>

39. Kong, et al., 2017. Spatial information literacy for digital Humanities: The case study of leveraging geospatial information for African American history education, College & Undergraduate Libraries, 24:2-4, 376-392, DOI: 10.1080/10691316.2017.1329044

40. Kulesz, O., 2011. Digital Publishing in Developing Countries: The Emergence of New Models?. Pub Res Q 27, 311–320. https://doi.org/10.1007/s12109-011-9241-4

41. Lievesley, D., 1995. Strategies for managing electronic archives. from http://www.ukoln.ac.uk/services/papers/bl/rdr6238/paper.html#lievesley

42. Limesurvey GmbH. / LimeSurvey: An Open Source survey tool /LimeSurvey GmbH, Hamburg, Germany. URL http://www.limesurvey.org

43. Masrek M. N., Khan A., 2016. Usability Assessment of Web Digital Library: A Survey Among Undergraduate Students. Journal of Information and Knowledge Management (JIKM) Volume 5 Number 2 (2015). pp 46-56

44. Meeks, E. and Grossner, K., 2012. ORBIS: An Interactive Scholarly Work On The Roman World. Journal of Digital Humanities, p.40.

45. Mergel et al., 2015. A method to support search string building in systematic literature reviews through visual text mining. In Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC '15). Association for Computing Machinery, New York, NY, USA, 1594–1601. DOI:https://doi.org/10.1145/2695664.2695902

46. Meyer, H. et al., 2015."Building Digital Archives: Design Decisions: A Best Practice Example." IEEE Conferences. N.p., 2015. 59–64. Web.

47. Muhammad, T., Khalid, M., & Farzana, S., 2010. Use of electronic information resources and facilities by humanities scholars. The Electronic Library. 28. 122-136. 10.1108/02640471011023423.

48. Murthy et al., 2007. Extending the 5S Digital Library (DL) Framework: From a Minimal DL towards a DL Reference Model.

49. Murthy, U. et al., 2010. "Extending the 5S Framework of Digital Libraries to support Complex Objects, Superimposed Information, and Content-Based Image Retrieval Services", Department of Computer Science, Virginia Polytechnic Institute & State University, p. Available at: https://vtechworks.lib.vt.edu/handle/10919/19380

50. Phiri, L. and Suleman, H., 2015. Managing Cultural Heritage: Information Systems Architecture. South Africa: Department of Computer Science, University of Cape Town, pp.1-13.

51. Ponto, J., 2015. Understanding and Evaluating Survey Research. Journal of the advanced practitioner in oncology, 6(2), 168–171.

52. Power et al., 2017. Improving Archaeologists' Online Archive Experiences Through User-Centred Design. J. Comput. Cult. Herit. 10, 1, Article 3 (April 2017), 20 pages. DOI:https://doi.org/10.1145/2983917

53. Raben, J., 1991. Humanities computing 25 years later. Computers and the Humanities. [Online] 25 (6), 341–350.

54. Ries, T. and Palkó, G., 2019. Born-digital archives. Int J Digit Humanities 1, 1–11 https://doi.org/10.1007/s42803-019-00011-x

55. Rockwell, G. and Mactavish, A., 2004. "Multimedia", A Companion to Digital Humanities, pp. 108-120. doi: 10.1002/9780470999875.ch10.

56. Rosa, C., 2020. Open source software for digital preservation repositories. International Journal of Computer Science & Engineering Survey (IJCSES) Vol.8, No.3, June 2017 Available at: https://arxiv.org/ftp/arxiv/papers/1707/1707.06336.pdf

57. Saltz, D., 1997. The Art Of Interaction: Interactivity, Performativity, And Computers. [ebook] Wiley on behalf of The American Society for Aesthetics, pp.The Journal of Aesthetics and Art Criticism, Vol. 55, No. 2, Perspectives on theArts and Technology (Spring, 1997), pp. 117-127. Available at: <https://www.jstor.org/stable/431258>

58. Sant, T., 2014. Interdisciplinary Approaches To Documenting Performance. [ebook] pp.International Journal of Performance Arts and Digital Media, 2014 Vol. 10, No. 1, 3–6,. Available at: <http://dx.doi.org/10.1080/14794713.2014.912495>

59. Sastry, H. and Reddy, L., 2009. User Interface Design Principles for Digital Libraries. International Journal of Web Applications.Volume 1 Number 2 . Available at: https://www.researchgate.net/publication/228680679_User_Interface_Design_Principles_for_Digital_Libraries

60. Schreibman, S., Siemens, R. and Unsworth, J., 2004. A Companion To Digital Humanities. Oxford: Blackwell.

61. Sinclair et al., 2013. "Information Visualization for Humanities Scholars". Literary Studies in the Digital. Available at: https://dlsanthology.mla.hcommons.org/information-visualization-for-Humanities-scholars/

62. Shen, R. et al., 2008. "Integration of complex archeology digital libraries: An ETANA-DL experience", Information Systems, 33(7-8), pp. 699-723. doi: 10.1016/j.is.2008.02.006.

63. Sherpa services /Welcome to OpenDOAR - v2.sherpa (2020). Available at: https://v2.sherpa.ac.uk/opendoar/

64. Slingerland, et al., 2017. The Distant Reading of Religious Texts: A “Big Data” Approach to Mind-Body Concepts in Early China, Journal of the American Academy of Religion, Volume 85, Issue 4, Pages 985–1016, https://doi.org/10.1093/jaarel/lfw090

65. Suleman, H., 2019. “Reflections on Design Principles for a Digital Repository in a Low Resource Environment.” CEUR Workshop Proceedings. Vol. 2461. CEUR-WS, 2019. 3–12. Print.

66. Svensson, P., 2010. The Landscape of Digital Humanities. Digital Humanities. 4:1

67. Terras, M., 2012. Infographic: Quantifying Digital Humanities | UCL UCLDH Blog. [online] Blogs.ucl.ac.uk. Available at: <https://blogs.ucl.ac.uk/dh/2012/01/20/infographic-quantifying-digital-Humanities/>

68. Terras, M., 2016. A Decade in Digital Humanities. Journal of Siberian Federal University. Humanities & Social Sciences. Vol 9. 1637-1650. DOI: 10.17516/1997-1370-2016-9-7-1637-1650.

69. The Consultative Committee for Space Data Systems., 2012 .Reference Model for an Open Archival Information System. Public.ccsds.org. Available at: https://public.ccsds.org/pubs/650x0m2.pdf

70. Thoden, Klaus, Juliane Stiller, Natasa Bulatovic, Hanna-Lena Meiners and Nadia Boukhelifa., 2017. “User-Centered Design Practices in Digital Humanities – Experiences from DARIAH and CENDARI.” ABI Technik 37 (2017): 11 - 2.

71. Thomas, C., 2017. Documents from the German Text Archive can now be analyzed with the Voyant Tools. Deutsches Textarchiv . Available at: http://www.deutschestextarchiv.de/news/61

72. Trenton et al., 2014. A Case Study for Universal Design in the Internet of Things. Assistive Technology Research Series. 35. 10.3233/978-1-61499-403-9-45.

73. Trimpi, W., 1970. The Definition And Practice Of Literary Studies. [ebook] The Johns Hopkins University Press, pp.New Literary History, Vol. 2, No. 1, A Symposium on Literary History (Autumn,1970), pp. 187-192. Available at: <https://www.jstor.org/stable/468596>

74. Urberg, M., 2017.Pasts and Futures of Digital Humanities in Musicology: Moving Towards a “Bigger Tent”, Music Reference Services Quarterly, 20:3-4, 134-150, DOI: 10.1080/10588167.2017.1404301

75. Unsworth, J., 2003. Tool-Time, Or 'Haven't We Been Here Already?' Ten Years In Humanities Computing. [online] People.virginia.edu. Available at: <http://www.people.virginia.edu/~jmu2m/carnegie-ninch.03.html>

76. Velmurugan, C. & Natarajan, R., 2015. Electronic Publishing: A Powerful Tool for Academic Institutions in the Electronic Environment. International Journal of Library Science and Information Management (IJLSIM). 1. 10-18.

77. VIIRI S., 2014. Digital Humanities and Future Archives. The Finnish Institute in London. United Kingdom, pp.1-40.

78. Wactlar, H.D. and Christel, M.G., 2002. Digital video archives: Managing through metadata. Building a national strategy for digital preservation: Issues in digital media archiving, 84, p.99.

79. Wasson, Christina, Gary Holton, and Heather Roth. Findings from the Workshop on User-Centered Design of Language Archives. N.p., 2016. Print.

80. Widlöcher, A., Lecarpentier, J., Mathet, Y., Bechet, N. and Roger, J., 2015. Combining Advanced Information Retrieval And Text-Mining For Digital Humanities. [ebook] New York, NY, USA: Association for Computing Machinery, pp.157–166. Available at: <https://doi.org/10.1145/2682571.2797067>

81. Zerkina, N. and Lomakina, Y., 2017. Linguistic And Digital Characteristicss Of Modern Information Environment. pp.Russian Linguistic Bulletin 2 (10) 2017.

82. Zhang, A., & Gourley, D., 2009. Selecting software and hardware for digital collections management systems. Creating Digital Collections, 149-177. doi: 10.1016/b978-1-84334-396-7.50010-9

Appendices

7.1. Appendix A: Survey Questions

Survey Question	Response options available to respondents	Justification for Survey Question
User Study Sample		
In general, do you consider yourself a Digital Humanist?	▪ Yes ▪ No	To ascertain whether Humanities researchers perceive themselves as users of digital resources, complex digital tools and information technologies in their scholarship.
Which one of the following best describes your occupation?	▪ General public ▪ High school student ▪ Undergraduate student ▪ Postgraduate student ▪ Postdoc researcher ▪ University academic ▪ Private sector professional ▪ Public sector professional ▪ Other	To determine the professional background of the Humanities researchers and to take note of the differences in their responses based on occupation.
Which of the following is/are your field(s) of Humanities?	▪ Anthropology ▪ Archaeology ▪ Classics ▪ History ▪ Linguistics and Languages ▪ Law and Politics ▪ Literature ▪ Philosophy ▪ Religion ▪ Performing Arts ▪ Instructional Technology ▪ Visual Arts ▪ Cultural Studies ▪ Library Sciences/Archives ▪ Media Studies/Media Arts ▪ Music ▪ Rhetoric and Writing ▪ Other:	To understand which of the fields of Humanities have the most representation in the survey responses.
Thinking about the Digital Humanities tools you use in your work/ teaching/ learning/ curriculum/ assessment/research, etc, please rate your technical proficiency and comfort with Digital Humanities tools.	▪ Highly proficient and very comfortable ▪ Proficient and comfortable ▪ Somewhat proficient and somewhat comfortable ▪ Not very proficient and uncomfortable ▪ Don't use digital tools	To assess where Humanities researchers are on their journey of acquiring the knowledge and skills necessary for using Digital Humanities tools and services.
I find technology useful as a tool for my work/ teaching/ learning/ curriculum/ assessment/research, etc.	▪ Strongly Agree ▪ Somewhat Agree ▪ Feel Neutral ▪ Somewhat Disagree	To predict the Humanities researcher's acceptance of information technology tools.

	▪ Strongly Disagree	
How often do you use digital archives?	▪ Daily ▪ Weekly ▪ Monthly ▪ More than once a year ▪ Once a year ▪ Less than once a year ▪ Never	To capture the usage frequency of digital archives by Humanities researchers and to establish a reasonable usage benchmark. To establish whether there are differences in the average usage of digital archives.
Which of the following methods, practices of teaching and objects of study used in Digital Humanities are you familiar with?	▪ Preservation of digital records ▪ Text Mining ▪ Data Visualization ▪ Spatial Analysis ▪ Image Analysis ▪ Digital Publishing ▪ Use of technology to study human objects and culture ▪ Use of traditional Humanities methods to study technology ▪ Use of digital archives to support student work in Humanities courses ▪ Student digital archiving of primary source materials ▪ Students participation in computational fieldwork, such as tagging data and creating metadata from primary source materials ▪ Students' use of blogs to track original undergraduate research and learning ▪ Student production of multimedia Humanities projects ▪ Other:	To measure the respondent's awareness and knowledge of services, methods, practices and objects of study used in Digital Humanities.
Access Resources		
If you think about the South Africa digital archives that you use regularly, how satisfied are you with the digital archive's process of making content easy to find and accessible to users? This is on a scale of 1 to 5, where 1 is satisfied and 5 is dissatisfied.	▪ 1 ▪ 2 ▪ 3 ▪ 4 ▪ 5	To get insights into whether services for making resources discoverable and accessible to researchers meet the expectations of Humanities researchers and if there are opportunities for optimisation.
Please rank the following functionalities found in digital archives in order of importance from 1 to 5, where 1 is most important to you and 5 is least important to you ▪ Browse - this function lists collections in a set order or organised according to a given characteristic or scheme.	▪ 1st option ▪ 2nd option ▪ 3rd option ▪ 4th option ▪ 5th option	To understand which functionalities of digital archives researchers like or value the most.

▪ Search – this function allows users to discover collections matching a query, which are then returned as a result set. ▪ Acquire – this function supports the user in retaining collections past the lifetime of the user's interaction with the system. ▪ Visualise – this function enables the user to view collections graphically. ▪ Create – this function supports users in creating new resources e.g. in creating new collections.		
Managing Resources		
Please evaluate each of the following statements. Please select one for each question: ▪ As a user of digital archives, I would like to be actively notified about updates to the digital collections. For example, notified every time collections have been updated or changed. ▪ As a user of digital archives, I would like to be allowed to make annotations/ comment on collections. ▪ As a user of digital archives, I would like to be allowed to create new digital objects by reusing existing digital objects, either in part or as a whole. For example, be allowed to compose a multimedia album by putting together audio files, song lists and singer biographies. ▪ As a user of digital archives, I would like digital archives to give me access to analysis services that allow me to examine digital objects. e.g. linguistic analysis, qualitative analysis services, statistical analysis services, scientific analysis services ▪ As a user of digital archives, I want to be able to perceive information or digital objects at different levels of detail, as desired.	▪ Strongly Agree ▪ Somewhat Agree ▪ Feel Neutral ▪ Somewhat Disagree ▪ Strongly Disagree	To understand the Humanities researcher's use of the functions that support the production, withdrawal or the update of digital resources.

▪ As a user of digital archives, I would like digital archives to give me access to conversion services that allow me to convert digital objects from one format to another e.g. Provide Text-to-speech, speech-to-text, spreadsheet-to-database, data-to-graphs, 3D-to-2D, and Image-to-colour histograms services. ▪ As a user of digital archives, I would like digital archives to provide translation services and multilingual information access ▪ As a user of digital archives, I would like digital collections to be extractable to removable media such as CD-ROM/DVD/USB flash drive. ▪ As a user of digital archives, I would like to see the quality status of the digital collections validated and the quality status displayed.		
Collaborate		
If Digital Archives in South Africa offered the following functionalities, how often do you think you would use the functions? Please select one for each function ▪ Share external communication details ▪ Communicate with other users through internal services ▪ Collaboratively author content	▪ Daily ▪ Weekly ▪ Monthly ▪ More than once a year ▪ Once a year ▪ Less than once a year ▪ Never	To capture the usage frequency of collaboration functions and to establish a reasonable usage benchmark. To establish whether there are differences in the average usage of collaborations functions.
Overall		
Overall, what has your experience been with South African Digital Archives?	▪ Very Satisfied ▪ Somewhat Satisfied ▪ Feel Neutral ▪ Somewhat Dissatisfied ▪ Very Dissatisfied	To evaluate the general sentiment of Humanities researchers towards South African digital archives, including their general level of satisfaction with the services provided by the archives. This is done to understand researchers overall experience.
If there was an additional functionality that you'd like added to digital archives and tools that you use in your work/ teaching/	Open for respondents' comments	To give respondents space to recommend and provide insight into useful functionality

learning/ curriculum/ assessment/research, what would it be?		that they would like to see available in digital archives.
What other, non-archive, digital tools do you use in your work/ teaching/ learning/ curriculum/ assessment/research, etc.?	Open for respondents' comments	To give respondents space to provide insights into other useful digital tools that they use regularly.
Closing		
Would you be willing to answer more questions if I need additional information in the future"? If so, please respond by filling in your contact information below:	Email address	To give respondents space to provide further information should it be required.
Would you like to receive a copy of the thesis at the completion of this research? If so, please respond by filling in your contact information below:	Email address	To get the respondents' contact information should they wish to receive a copy of the thesis.

7.2. Appendix B: Informed Consent Form

DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF CAPE TOWN
PRIVATE BAG X3
RONDEBOSCH 7701
SOUTH AFRICA

RESEARCHER/S: Khanyisa Mtombeni
TELEPHONE: +27799827844
E-MAIL: MTMKHA001@myuct.ac.za
URL: https://www.cs.uct.ac.za/

Informed Voluntary Consent to Participate in Research Study

Project Title: Investigating how South African Humanities Researchers Engage with Digital Archives

Invitation to participate, and benefits: You are invited to participate in a research study conducted with Humanities scholars. The study aim is to investigate how South African Humanities researchers engage with Digital Archives. I believe that your experience would be a valuable source of information, and hope that by participating you may gain useful knowledge.

Procedures: During this study, you will be asked to provide responses to a research survey based on your experience using digital archives.

Risks: There are no potentially harmful risks related to your participation in this study.

Feedback: You will receive feedback about the results of this research in the form of a thank you email with a copy of the research report attached.

Disclaimer/Withdrawal: Your participation is completely voluntary; you may refuse to participate, and you may withdraw at any time without having to state a reason and without any prejudice or penalty against you. Should you choose to withdraw, the researcher commits not to use any of the information you have provided without your signed consent. Note that the researcher may also withdraw you from the study at any time.

Confidentiality: All information collected in this study will be kept private in that you will not be identified by name or by affiliation to an institution. Confidentiality and anonymity will be maintained as pseudonyms will be used.

What signing this form means: By signing this consent form, you agree to participate in this research study. The aim, procedures to be used, as well as the potential risks and benefits of your participation have been explained verbally to you in detail, using this form. Refusal to participate in or withdrawal from this study at any time will have no effect on you in any way. You are free to contact me, to ask questions or request further information, at any time during this research.

I agree to participate in this research (tick one box) ☐ Yes ☐ No _____ (Initials)

_____	_____	_____
Name of Participant	Signature of Participant	Date
_____	_____	_____
Name of Researcher	Signature of Researcher	Date

7.3. Appendix C: Research Ethics Clearance

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Faculty of Science
University of Cape Town
Rondebosch
South Africa 7701

E-mail: shari.daya@uct.ac.za
Tel: 021 650-2880

7 April 2020

Mr Khanyisa Mtombeni
Department of Computer Sciences

Investigating how South African Humanities Researchers Engage with Digital Archives

Dear Mr Khanyisa Mtombeni

I am pleased to inform you that the Faculty of Science Research Ethics Committee has approved the above-named application for research ethics clearance, subject to the conditions listed below.

- Restrictions on involving human participants in research must be adhered to, given current concerns about the spread of Covid-19. Please ensure that you are aware of and comply with UCT policy on this, as communicated by management.
- Please also, obtain the requisite permission/s to conduct research with UCT staff and students.
- Implement the measures described in your application to ensure that the process of your research is ethically sound; and
- Uphold ethical principles throughout all stages of the research, responding appropriately to unanticipated issues: please contact me if you need advice on ethical issues that arise.

Your approval code is: **FSREC 016- 2020**

I wish you success in your research.

Yours sincerely

Signed by candidate

Dr Shari Daya
Chair: Faculty of Science Research Ethics Committee

Cc: A/Prof Hussein Suleman (supervisor)