Metadata schema for oral history content in South Africa

by

MATLALA RACHEL MAHLATJI

30183162

Submitted in accordance with the requirements for the degree of

DOCTOR OF PHILOSOPHY

in the subject

Information Science

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: Prof. M Ngoepe

2023

ABSTRACT

Metadata schemas are becoming increasingly important, not only for storing internal data, but also for allowing content exchange, reuse, and access. Metadata are also important for the retrieval of oral history content. In the dawn of democracy in South Africa, oral histories of previously marginalised communities were collected to close the gaps in archival holdings. The National Film, Video and Sound Archives (NFVSA) was charged with the responsibility of keeping the collected content in its custody. The NFVSA was mandated to make oral history content available and accessible through metadata creation. However, the creation of metadata for oral history content presents some challenges that affect its accessibility. This study explored the metadata schema for oral history content stored in the NFVSA, with the goal of proposing a framework to ensure access and retrievability of oral history content. Qualitative data were collected through document analysis, inventories of oral history output at the NFVSA, and interviews with purposively selected archivists at the NFVSA charged with the curation of audio-visual archives. Data was thematically analysed. The findings suggest that oral history content collected over years since the dawn of democracy in South Africa is inaccessible because it is uncatalogued or not classified. As a result, the outputs collected and converted to audio-visual formats were stored in boxes. This content may become irretrievable in future due to technological obsoleteness or it has not been arranged and described. It was concluded that failure to transform this pattern would result in the loss of collected oral content that was supposed to be used for the transformation of the archive holding. It is recommended that the NFVSA should create metadata to ensure that oral history content can be retrieved and accessed. This study proposes a metadata schema and an access framework for oral history content in South Africa. The framework may be applicable to other repositories in the country and on the continent. A further study on the selection of oral history content to be captured is suggested, as the current collection still reflects the current elite members of society, thereby perpetuating elitism in critical emancipatory.

Keywords: data at risk, digitisation, FAIR, metadata, metadata creation, metadata schema, metadata standards, metadata system, migration, oral history content

ACKNOWLEDGEMENTS

First and foremost, I thank Almighty God for His grace, strength, and guidance throughout this academic journey.

I extend my sincere gratitude to **Prof. M. Ngoepe** for his invaluable supervision, advice, encouragement, and unwavering support throughout this study.

My appreciation goes to the **University of South Africa (UNISA)**, particularly the Department of Information Science in the School of Arts, College of Human Sciences, for providing me with the opportunity to undertake this research.

I am grateful to the **National Film, Video and Sound Archives (NFVSA)** for granting permission to conduct this study and to all staff members who contributed their time, knowledge, and support.

Special thanks go to **Ms Morongoe Ntsala** for her professional guidance and motivation. I also thank my academic friend, **Prof. Maoka Dikotla**, for his support, motivation, and scholarly engagement throughout this journey.

I am deeply grateful to my mother, **Mrs Anna Mosholadi Mahlatji**, for her unconditional love, prayers, sacrifices, and unwavering belief in me. To my beloved children, **Kagisho, Kholofelo, Mosholadi, and Mamphere**, thank you for your patience, understanding, and constant encouragement.

I also thank **Information Giants (Pty) Ltd** and **Ms Letitia Greenberg** for their professional editing and valuable suggestions.

Finally, I extend my appreciation to my family, colleagues, friends, and everyone who supported and encouraged me throughout this study.

This achievement reflects the collective support, encouragement, and contributions of all those who walked this journey with me, and I am sincerely grateful to each one of you.

DEDICATION

I dedicate this work to my eldest son, **Kagiso David,** and his siblings, **Kholofelo, Mosholadi Annette,** and **Mampheri Nociah**; my mother, **Mosholadi Anna;** and my late father, **Mashabela Joseph Mahlatji.**

DECLARATION

Name: Matlala Rachel Mahlatji

Student number: 30183162

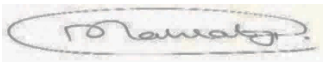
Degree: Doctor of Philosophy in Information Science

Metadata schema for oral history content in South Africa

I declare that this these **Metadata schema for oral history content in South Africa** is my own original work and that all the sources I have cited and quoted have been indicated and referenced fully as required by the university's plagiarism policy.

I further declare that I submitted this thesis to originality checking software and that it falls within the accepted requirements for originality.

I further declare that I have not previously submitted this work, or part of it, for examination at Unisa for another qualification or at any other higher institution of education.



.....

Name: Mahlatji, Matlala Rachel Date: 14 May 2024

TABLE OF CONTENTS

Abstract i
Acknowledgements ii
Dedication iii
Declaration iv
TABLE OF CONTENTS i
List of abbreviations and acronyms vii
List of figures viii
List of tables ix
Chapter One 1
Introduction: Setting the Scene 1
1.1 Introduction and background to the study 1
1.1.1 Contextual setting 3
1.2 Problem statement 9
1.3 Purpose and objectives of the study 10
1.4 Conceptual framework 10
1.4.1 The Metadata Objectives and Principles, Domains, and Architectural Layout 12
1.4.2 The Theory of Metadata Enriching and Filtering 12
1.4.3 The Open Archives Information System framework 14
1.4.5 Open Archives Initiative Protocol for Metadata Harvesting 17
1.4.7 International Standards Organisation 17
1.4.8 Policies and guidelines in archives 19
1.4.9 Oral History Association and oral history best practises 20
1.5 Originality of the study 21
1.6 Justification of the study 22
1.7 Scope and delimitation of the study 23
1.8 Literature review 24
1.9 Research methodology 25
1.10 Dissemination of the findings 26
1. 11 Ethical Considerations 26
1.12 Definition of Concepts 27
1.12.1 Classification schema 27
1.12.2 Fonds 28

11.2.3 Metadata....28
1.12.4 Metadata schema....**Error! Bookmark not defined.**
1.12.5 Metadata standard....28
1.12.6 Oral history....28
1.12.7 Oral history content....**Error! Bookmark not defined.**
1.12.8 Access to oral history....29
1.13 Structure of the thesis....29
1.14 Summary....30
Chapter Two....31
Literature Review....31
2.1 Introduction....31
2.2 Purpose of the literature review....**Error! Bookmark not defined.**
2.3 Literature map for this study....31
2.4 ORAL HISTORY CONTENT in archival institutions....33
2.5 Metadata schema for oral history content....34
2.5.1 Types of metadata....34
2.5.2 Metadata elements....37
2.5.3 Metadata schema in archival institutions....38
2.5.4 Metadata quality and metadata creation....40
2.5.5 Selection of Metadata schema....43
2.6 Metadata standards for oral history content....44
2.6.1 Principles of metadata standards....44
2.6.2 Types of metadata standards....47
2.6.3 Classification of oral history content....48
2.6.4 Metadata standards and oral history content....49
2.7 Access to oral history content....51
2.7.1 Formats of oral history content....52
2.7.2 Components on access to oral history content....52
2.8 Technologies used for metadata of oral history content....55
2.8.1 Information management systems in archival institutions....55
2.8.2 Databases and finding aids in archival institutions....58
2.8.3 Software technologies in archival institutions....59
2.8.4 Artificial intelligence and metadata creation....61
2.8.5 Digitisation of oral history content....66

2.8.6 Migration of metadata for oral history 67
2.9 Framework for metadata for oral history content 69
2.10 Summary 71
Chapter Three 72
Research Methodology 72
3.1 Introduction 72
3.2 Research paradigm 73
3.2.1 Epistemology 74
3.2.2 Ontology 78
3.2.3 Choice of the paradigm 81
3.3 Research approach 83
3.3.1 Quantitative approach 83
3.3.2 Qualitative approach 84
3.3.3 Mixed methods approach 85
3.3.4 Choice of the approach 86
3.4 Research design 86
3.4.1 Research designs in qualitative research 87
3.4.2 Choice of research design 89
3.5 Population and sampling 90
3.5.1 Population 90
3.5.2 Sample and sample procedure 92
3.5.3 Sample size 94
3.6 Data collection techniques 94
3.6.1 Document analysis 94
3.6.2 Interviews 95
3.6.3 Data collection procedure 98
3.7 Data analysis 101
3.8 Research rigour 102
3.8.1 Reliability and validity 102
3.8.2 Credibility, transferability, and conformability 105
3.9 Ethical Considerations 105
3.9.1 Permission to conduct the study 106
3.9.2 Informed consent 107
3.9.3 Anonymity and confidentiality 108

3.9.4 Scientific integrity 109
3.10 Evaluation of research methodology 111
3.11 Summary 112
Chapter Four 113
Data Analysis and Presentation 113
4.1 Introduction 113
4.2 Data analysis and presentation 113
4.2.1 Study participants 114
4.2.2 Data analysis 115
4.3 Data presentation 117
4.3.1 Metadata schema used for oral history content 117
4.3.2 Metadata standards used for the classification of oral history content 124
4.3.3 Access to oral history content at the NFVSA 130
4.3.4 Technologies used for the metadata and access to oral history content 135
4.3.5 Metadata schema framework for oral history content 141
4.4 Themes from data analysis 147
4.5 Summary 148
Chapter Five 149
Data Interpretation and Discussion 149
5.1 Introduction 149
5.2 Metadata schemas used for oral history at the NFVSA 150
5.2.1 Knowledge of metadata schema 151
5.2.2 Metadata schema used at the NFVSA 152
5.2.3 Metadata elements for oral history content 153
5.2.4 Metadata creation process 153
5.2.5 Metadata categories recorded 153
5.3 Metadata standards for oral history content at the NFVSA 154
5.3.1 Knowledge of metadata standards 155
5.3.2 Classification standards for oral history content 155
5.3.3 Classification guidelines for oral history content 156
5.3.4 Choice of metadata standards 156
5.4 Access to oral history content at the NFVSA 156
5.4.1 Formats of oral history content 157
5.4.2 Search strategies to access oral history content 157

5.4.3 Ethics and legal access to oral history content....157
5.5 Technologies used for metadata schema of oral history content at the NFVSA....158
5.5.1 Technologies for access to oral history content....159
5.5.2 Finding aids and access to oral history content....159
5.5.3 Technologies for metadata creation....160
5.5.4 Benefits of technologies used for oral history content....160
5.6 Proposed metadata schema framework for oral history content....162
5.7 Summary....166
CHAPTER Six....168
Summary, conclusions, and recommendations....168
6.1 Introduction....168
6.2 Summary of the findings....168
6.2.1 Metadata schemas used for oral history content....169
6.2.2 Metadata standards for oral history content....169
6.2.3 Access to oral history content....170
6.2.4 Technologies used for oral history content....171
6.2.5 Proposed metadata schema framework for oral history content....172
6.3 Conclusions from the findings of the study....173
6.3.1 Metadata schema used for oral history content at the NFVSA....174
6.3.2 Metadata standards for oral history content....174
6.3.3 Access to oral history content in the NFVSA....175
6.3.4 Technologies used for oral history content at the NFVSA....176
6.3.5 Proposed metadata schema framework for oral history content....176
6.4 Recommendations....177
6.4.1 Metadata schema of oral history content....177
6.4.2 Metadata standards for oral history content....177
6.4.3 Access to oral history content....178
6.4.4 Technologies used for oral history content....178
6.5 Proposed metadata schema framework for oral history content....179
6.6 Implications of theory and scholarship....188
6.7 Recommendations for further research....190
6.7.1 Information search patterns for the archival institutions....191
6.7.2 Collaborative approach to oral history collections in South Africa....191
6.7.3 Oral history metadata schema in academic institutions....192

6.7.4 Legal framework for oral history content 192
6.7.5 Digital transformation in archival services 192
6.7.6 Transforming metadata schema system in archival institutions 193
6.7.7 Digitisation in archival institutions 193
6.7.8 Database searches and access to archival materials 194
6.7.9 Decolonisation and transformation in archival institutions 194
6.8 Final conclusion 195
References 196
Appendix A: Participation Information Sheet 225
Appendix B: Informed Consent 230
Appendix C: Ethical clearance letter 231
Appendix D: Data Collection Tools 233

LIST OF ABBREVIATIONS AND ACRONYMS

AtoM	Access to Memory
DAC	Department of Arts and Culture
ICA	International Council on Archives
ISAAR	International Standard Archival Authority Record
ISAD (G)	International Standard for Archival Description (General)
ISDIAH	International Standard for Describing Institutions with Archival Holdings
ISO	International Organisation for Standardisation
LOM	Learning Object Metadata
NAAIRS	National Automated Archival Information Retrieval System
NAROS	National Register for Oral Sources
NARSA	National Archives of South Africa Act
NARSSA	National Archives and Records Services of South Africa
NFVSA	National Film, Video and Sound Archives
OHA	Oral History Association
OHASA	Oral History Association of South Africa
PAIA	Promotion of Access to Information Act
PAJA	Promotion of Administrative Justice Act
POPIA	Protection of Personal Information Act
SAA	Society of American Archivists

List of figures

Figure 1.1: South Africa on the World Map (https://southafricamap360.com/south-africa-world-map).....4

Figure 1.2: Nine Provinces of South Africa (https://www.mappr.co/counties/south-africa/) ...5

Figure 1.3: Map for the location of Pretoria (www.Gauteng-Map.asp) 5

Figure 1.4: Conceptual framework for the study (Researcher's own work).....11

Figure 1.5: The Theory of Metadata Enriching & Filtering (Alemu 2015).....13

Figure 2.1: Literature Map.....32

Figure 3.1: Research map (..... **Error! Bookmark not defined.**

Figure 4.1: Search results for Sharpeville Massacre survivors.....133

Figure 4.2: NAAIRS Search function: Old and/or new database 137

Figure 4.3: Inventories for oral history content 138

Figure 4.4: Contributions to translations 146

Figure 4.5: Summary of Themes (Researcher own work 147

Figure 6.1: Framework for metadata and access to oral history in archival institutions (Researcher' own work).....180

LIST OF TABLES

Table 3.1 Research Paradigm

Table 4.1: Breakdown of staff at the NFVSA.....114

CHAPTER ONE
INTRODUCTION: SETTING THE SCENE

1.1 INTRODUCTION AND BACKGROUND TO THE STUDY

Metadata schemas are valuable tools for creating and managing metadata in organisations as well as facilitating metadata facilitate the discovery, use, preservation, and sharing of information resources. Eíto-Brun (2015) notes that metadata provide comprehensive descriptions of resources that support discovery, access, preservation, and sharing. Similarly, Kim (2020) explains that metadata describe the who, what, when, where, why, and how of a dataset or resource. Fraser and Halvey (2017) further emphasise that metadata are essential for describing, organising, and sharing information, including oral history content. Since oral history content must be accessible for use and reuse, effective description is critical. Tembe (2019) argues that access is a central component of preservation, highlighting the importance of ensuring that oral history content remains accessible to society.

Although several studies (Ritchie 2014; MacKay 2016; Bradley & Puri 2018; McDonald, Puddicombe, Weymouth & Williams 2021; Thor Tureby & Wagrell 2022) have focused on collecting and curating oral history, limited attention has been given to access after collection. Oral history content plays a significant role in addressing historical distortions and amplifying marginalised voices. It can form part of historical records, support cultural heritage preservation, serve as legal evidence, contribute to genealogical research, verify authenticity, and satisfy personal interests (Pryse, Harp, Mannheimer, Marsolek & Cowles 2021). Moss and Mazikana (1986) view oral history as part of society's heritage that helps restore history and strengthen culture. Likewise, Ngoepe (2019) argues that oral history is essential for understanding the past of communities whose histories were never documented in written form. However, oral history content in South Africa often remains inaccessible because it is uncatalogued, limiting intellectual control and access. Most of this content is preserved in public archives repositories.

Public archives are mandated to provide access to the nation's archival heritage, including oral history content. Ngulube, Sibanda and Makoni (2013) identify access to primary archival data as a core function of archival administration, while Venson, Ngoepe and Ngulube (2014) recognise archives as sources of demographic, economic, and social development. Ngoepe

(2019:159) further argues that citizens require access to archives and records to understand the past, hold governments accountable, and address historical inequalities. Achieving such access depends on the availability of high-quality metadata, making metadata schemas essential for organising and providing access to oral history content.

A metadata schema is a system of labelling, tagging, and coding used to describe and structure information resources (Xie & Matusiak 2016). Metadata schemas are increasingly important for information management, sharing, and reuse (Dashrath 2014). Although no single schema is suitable for all resources, archival institutions often adopt widely recognised standards. For example, Encoded Archival Description (EAD) is a standard developed by the Society of American Archivists for encoding archival finding aids (Gracy & Lambert 2014). Institutions must therefore adapt metadata schemas and standards to improve accessibility and service delivery.

The use of metadata schemas significantly influences the organisation and accessibility of archival collections. Archivists seek to implement schemas that promote interoperability and information sharing across archives (Da Graça Simões, De Freitas & Rodríguez-Bravo 2016). According to Miller (2022), factors influencing schema selection include user needs, access requirements, relationships with other collections, collection size and scope, interoperability, and maintenance requirements. Turner (2017) adds that the nature of the information being described should guide the choice of metadata schema. Consequently, archival institutions should adopt schemas that are appropriate for their oral history collections.

Oral history content may consist of digitised analogue audio and video recordings or materials captured using mobile phones, audio recorders, video recorders, and other digital devices (Pryse et al. 2021). Mosweu (2011) notes a shift from analogue recording technologies to digital equipment for creating oral history records. Hendrigan (2016) further highlights the importance of preserving and providing access to oral history collections as audio-visual resources in South Africa, making information and communication technologies (ICTs) an important consideration for archival institutions.

Despite the availability of metadata standards and schemas, many oral histories remain uncatalogued and inaccessible (Wynne 2009). Ncala (2017) found that a lack of cataloguing contributed significantly to archival inaccessibility, while Ngoepe (2019) observed that

classification is often overlooked despite its importance. Similarly, Ngoasheng, Ngoepe and Marutha (2021) found that many South African Broadcasting Corporation (SABC) audio-visual archives in Limpopo were difficult to access because they were uncatalogued. Existing standards and schemas may also be inadequate for capturing all the information required to describe oral history collections. Consequently, there is a need to develop guidelines or frameworks specifically suited to oral history metadata. Rawson (2018) argues that metadata are essential for ensuring that resources remain discoverable, accessible, and useful over time.

Metadata are therefore critical for enabling the discovery, access, and sharing of digital resources (Park & Tosaka 2010:104). However, the National Film, Video and Sound Archives (NFVSA), the custodian of South Africa's oral history collections, appears to face metadata-related challenges. Many oral history records lack detailed descriptive metadata, and some submissions are incomplete, requiring archivists to obtain missing information from contributors. Although these gaps may have limited effects on general users, they can affect research value and long-term accessibility. Feliciati and Alfier (2014) argue that improved access can be achieved through metadata schemas that adequately describe oral history content. Accordingly, this study investigated metadata schemas for oral history content in South Africa, with particular focus on the NFVSA and its audio-visual oral history collections.

1.1.1 Contextual setting

Public archives of South Africa are tasked with preserving a nation's archival heritage for the government and the public. This includes acquiring, maintaining and making accessible records of national significance, including both public and non-public records. South Africa is located at the southernmost point of the Africa continent, bordering both the Indian and the South Atlantic oceans (See Figure1.1).



Figure 1.1: South Africa on the World Map

(https://southafricamap360.com/south-africa-world-map)

South Africa had four provinces prior to 1994: the Cape, Natal, Transvaal, and Orange Free State. With the adoption of a new democratic constitution in 1996, the homelands were dismantled and South Africa as a country was divided into its nine provinces as it exists today. The nine provinces are: the Eastern Cape, the Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, the Northern Cape, the Northwest, and the Western Cape, as shown in Figure 1.2.



Figure 1.2: Nine Provinces of South Africa (https://www.mappr.co/counties/south-africa/)

Due to its many different cultures and religions, South Africa was dubbed the "rainbow nation." To name a few, Zulu, Xhosa, Sotho, Tswana, Ndebele, Khoisan, Hindu, Muslim, and Afrikaner people live within South Africa's borders (South African History Online (SAHO) 2019). These cultures require the collection of oral history content to help identify, appreciate, and respect various aspects of heritage. Oral history is crucial in 'bringing together disparate strands of South Africa's heritage", (Oelefse & Du Bruyn 2004: 156). This heritage will be enhanced by the contributions of each province.

The National Archives and Records Service (NARSA) is based in Pretoria, Gauteng province (See Figure 1.3 for a map of the Gauteng province showing the location of Pretoria). Pretoria is not only South Africa's administrative capital, but is also a city of government and culture, with theatres, museums, monuments, and tree-lined avenues. The NFVSA's headquarters are in Pretoria.



Figure 1.3: Map for the location of Pretoria (www.Gauteng-Map.asp)

The National Archives and Records Service of South Africa (NARSSA) is the 'central archival repository' in the Republic of South Africa (Molefe & Schellnack-Kelly 2019). The NFVSA is the country's only national institution with a one-of-a-kind mandate because it is the only archive mandated with the collection of, and access to, audio-visual content, including oral

history content. This section discusses the NFVSA's mandate and the oral history collection in South African archives to provide the context of this study.

1.1.1.1 National Film, Video and Sound Archive

National Archives and Records Services of South Africa (NARSSA) was established in 1996 by the National Archives and Records Service Act (Act No. 43 of 1996), as amended (NARSSA Act of 1996). South Africa has 10 public archives repositories: nine provincial archives and one national archive, with the NFVSA as its subsidiary (Ngoepe 2019). The NFVSA is a government institution established under the NARSSA Act and is in charge of preserving the country's audio-visual heritage for current and future generations. At the NFVSA, where citizens can access audio-visual material, various formats such as sound recordings, and other materials such as scripts, posters, and related are found. The NFVSA serves as a resource centre for South Africa's audio-visual heritage (NARSSA Act 1996) and as a legal deposit facility for film, video, and sound archives.

Mukwevho and Ngulube (2022:2) point out that NARSSA and the NFVSA, its subsidiary, are mandated by law to organise, describe, preserve, and make public records accessible to members of the public. The audio-visual collections at the NFVSA, together with the national and provincial archival services, 'archival heritage of South Africa spreads from the diaries of Jan van Riebeeck to oral history projects with communities in and around South Africa', (Zimu-Biyela 2022:4). The NARSSA Act includes a provision for the legal deposit of audio-visual material at the request of the national archivist. The Legal Deposit Act (Act No. 54 of 1997) designates the NFVSA as a deposit location for audio-visual materials published and made available in South Africa, and all provinces in South Africa are required to deposit their metadata in the National Automated Archival Information Retrieval System (NAAIRS), which receives contributions from the provincial archives. In addition to the audio-visual collection, the NFVSA has slides, posters, scripts, a library, and photographs dealing with audio-visual material, as well as information on the arts such as literature, music, sculpting, painting, and theatre (NARSSA 1996). Given that some of these audio-visual records have been stored at the NFVSA for more than 50 years (1964-2017), their physical carriers are becoming increasingly frail (Ncala 2017). Ngoepe (2019) also mentions that some of the oral history collection is kept in boxes in the form of tapes and compact discs. As a result, there are no metadata for these contents; as such, they are inaccessible to the general public.

The NFVSA has a broad mandate to collect audio-visual and related material made in or about South Africa. Most of the materials are donated by the film, video, and sound industries, as well as private individuals. The NFVSA also preserves the country's audio-visual heritage through oral history content.

1.1.1.2 Oral history content in South Africa

Oral history refers to a purposeful eliciting and preserving spoken memories of events and experiences of a person based on first-hand information, typically in audio or audio and visual recording media, is known as oral history (Boyd 2011). Because oral history preserves the perspectives and experiences of people and groups that might be underrepresented in written works, it can aid in closing gaps in the historical record (Schellnack-Kelly 2017). Oral history content can help people remember their past and, in some cases, resolve unpleasant memories of their past. As Ngoepe (2022) states:

> The oral source is an important part of recreation and rethinking the past, especially for those who never had an opportunity in the past. While archivists look to the past and talk to the future, records of the current government have not been properly managed, and therefore the public sector, especially municipalities, are characterised by disclaimer opinions, year after year.

To fill the gaps in oral history content in archival institutions, the South African government established the National Oral History Programme (NOHP) in 2001 and the Oral History Association of South Africa (OHASA) in 2003. The White Paper on Arts and Culture outlines government policy for the promotion, protection, creation, and funding of South African arts, including written and oral literature, culture, heritage, and related practitioners. As a result, the NOHP is consistent with this White Paper. The right to a culture is explicitly stated in the 1996 Constitution [Sections 30 and 31]. The Department of Arts and Culture (DAC) was tasked with conceptualising and spearheading South Africa's NOHP, which aims to produce information that will be added to what is already available in the country's archival holdings. This NOHP is also critical to the successful implementation of sustainable development. The NOHP began with a pilot project titled "The 1956 Anti-Pass March to the Union Buildings by South African Women." This topic was chosen because it is one of under-documented historic events. At that point, NARSSA, among others, has also collected stories of forced removals of communities

in Marabastad and forced relocation of people to make way for game parks like the Kruger National Park.

Molobye (2014) laments that in South Africa, despite their existence and involvement in oral history, community archives like the Royal Bafokeng Archives are not a part of the official national archive system. The National Registry of Oral history projects does not even contain these projects (Ngoepe 2019). In South Africa, OHASA seeks to support and enable oral history research, recording, preservation, accessibility, and popularisation. Poetry, music, oral tradition, oral acclaim, and oral performance all fall under this category. The mission of OHASA is to promote and facilitate oral history recording, preservation, access, popularisation, and study in South Africa. NARSSA continues to support OHASA in hosting national annual oral history conferences in all provinces. NARSSA's support for OHASA is primarily intended to allow practitioners to engage in and debate current trends in oral history theory and methodology. The purpose of the conferences is to promote and facilitate the recording, preservation, access, and popularisation of oral history in South Africa (including poetry, music, oral praise, oral performance, and oral traditions). The hosting of the annual conferences provided a platform for organic intellectuals to share their knowledge with delegates and practitioners. Bhebhe and Ngoepe (2022) commend OHASA for its role in the coordination of oral history programmes in South Africa. Access becomes more visible because of this success.

The DAC was aware of the need for oral history training for staff from the DAC and NARSSA. NARSSA, in collaboration with the provincial archives and OHASA, trained students on oral history methodology and how to use it to research or build a family tree (NARSSA 2016). Through the recording and archiving of oral histories, the DAC has the opportunity to record the largely untold micro-histories of South Africans. The practice of recording oral histories has two major benefits: (1) acting as a transformational mechanism for addressing past imbalances and functioning as a therapeutic mechanism for dealing with South Africa's turbulent past and (2) enriching our educational curriculum at all levels by creating deep and thorough archives accessible to domestic and international scholars (Zungu, Manqele, De Vries, Molefe & Hadebe 2014).

NARSSA is also in charge of maintaining a National Register of Oral Sources (NAROS), where oral stories are registered. This database is managed to provide interested parties with leads on research being conducted on subjects of interest. All oral history collections preserved

at the National Archives are preserved to enable access. For example, NARSSA collected stories from eyewitnesses who had fought and witnessed the struggle against the introduction of the Homeland System, as well as against land dispossession and the imposition of Bantustan leaders, with the Ingquza Hill Peasant Revolt. These collections are available to researchers, students, and community members, and are used for research purposes. Scholars and practitioners use them for a variety of purposes, including short story writing, drama, filmmaking, and historical research.

1.2 PROBLEM STATEMENT

Archival institutions require metadata schemas to facilitate information retrieval and/or improve accessibility of oral history content. The metadata schema's description of oral history content ensures interoperability among archives while increasing usage and awareness of these sources and services (Lee 2017). Currently, archival institutions use the International Standard Archival Description (General) (ISAD(G) metadata standard. ISAD(G) is an international standard that specifies the content of an archival description (ISADG). Oral history recordings and collections are frequently incompatible with existing standards developed for other types of materials and collections (Lee 2017). Furthermore, the documentation of oral history recordings contains numerous inaccuracies due to the problematic manner in which it was recorded (Ngoepe 2019). As a result, "oral history collections are often underused or tucked away in a dark archive, accessible only to researchers, as is the case with the NFVSA" (Ngoepe 2019:158). Bosum and Dunne (2017) concur that oral history content often lurks in the shadows, making access difficult. Tembe (2019) states that ineffective preservation and access strategies for oral history records resulted in ineffective preservation and access. According to Fraser and Halvey (2017), the lack of standardisation limits access to oral history content due to inconsistent metadata fields across collections. This has an impact on the accessibility of oral history content. If oral history content is not accessible, African knowledge and many narratives may be lost (Ngoepe 2020). Given that access and usage are the primary goals of an archival institution, it is critical to pay close attention to the metadata schema to provide awareness and access to its users; therefore, it is necessary to investigate the metadata schema for oral history content in South Africa. Hence, the purpose of this study was to propose a metadata schema framework for oral history collections in South Africa.

1.3 PURPOSE AND OBJECTIVES OF THE STUDY

This study explored the metadata schema for oral history content in South Africa, with the aim of proposing a framework to ensure its accessibility and retrievability. The specific objectives were to:

- describe the metadata schema used for oral history at the NFVSA
- analyse metadata classification standards for oral history content at the NFVSA
- determine access to oral history content at the NFVSA
- determine the technologies used for metadata of oral history content at the NFVSA
- propose a metadata schema framework for oral history content.

1.4 CONCEPTUAL FRAMEWORK

A conceptual framework provides a structure for understanding relationships among the key concepts that underpin a study. In qualitative research, conceptual frameworks are particularly useful because they enable researchers to draw on multiple theories and concepts to explain a complex phenomenon (Crawford 2019; Ngulube 2020).The present study explores metadata schema for oral history content, a topic that intersects archives, information science, technology and access to information. As no single theory sufficiently explains all aspects of metadata creation, management and retrieval, a conceptual framework was developed by integrating concepts drawn from metadata theory, archival standards, oral history practice and digital preservation literature.

The conceptual framework brings together six key concepts, namely metadata schema, metadata standards, metadata creation, technologies supporting metadata, user access and oral history content. These concepts are linked by the assumption that the quality and structure of metadata influence the accessibility and retrievability of oral history resources. Figure 1.4 illustrates the interconnection between these concepts and how metadata and access to oral history concepts relate to each other and support each other to form a conceptual framework for this study.

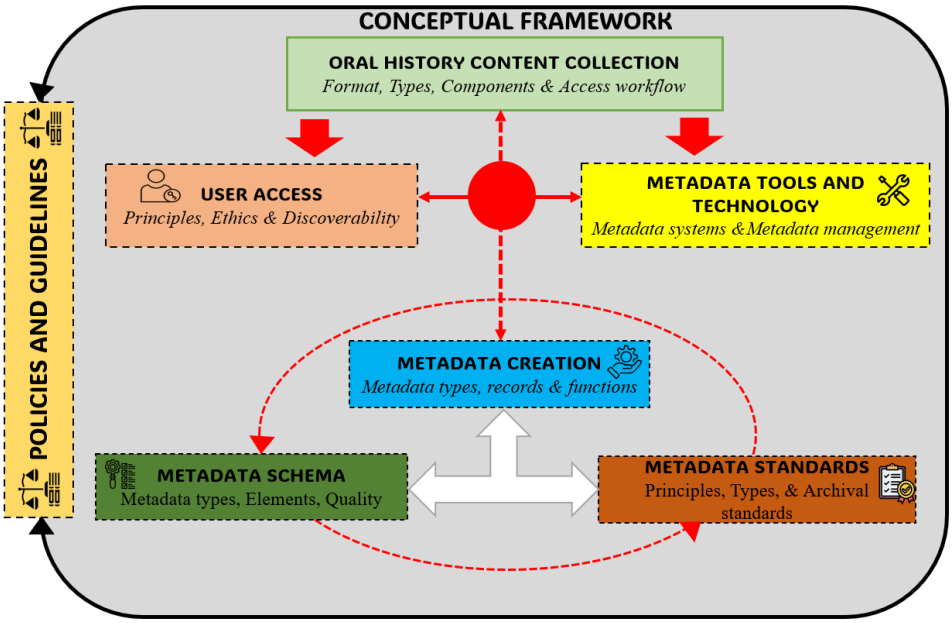


Figure 1.4: Conceptual framework for the study (Researcher's own work)

The conceptual framework for the current study was developed using concepts from the study objectives, specifically metadata schema, metadata standards, metadata creation, metadata tools and technology, user access, and oral history content. The conceptual framework connected the title, problem statement, and objectives of this study. However, it can be difficult to identify just one theory in the field of library and information studies (LIS). Theories from other disciplines, such as education, psychology, sociology, and information technology, are frequently used, but they are not LIS theories. As a result, the researcher had to employ multiple theories to develop a conceptual framework for this study that reflected all the objectives.

The following theories and models informed the conceptual framework for this study: Theory of Metadata Enrichment and Filtering, Metadata Objectives and Principles, Domains, and Architectural Layout (Modal), The Open Archives Information System framework, Open Archives Initiative Protocol for Metadata Harvesting, Oral History Association (OHA) best practices, and the International Standards for Organisation. Furthermore, the legislative framework governing metadata and access to oral history content was used to guide the application of the theories underlying this study.

1.4.1 The Metadata Objectives and Principles, Domains, and Architectural Layout

The Metadata Objectives and Principles, Domains, and Architectural Layout (MODAL) framework, developed by Greenberg (2005), provides an approach for analysing metadata schemas. According to Greenberg (2005), metadata schemas can be understood by examining three key components: objectives and principles, domains, and architectural layout.

Objectives and principles refer to the goals a metadata schema seeks to achieve and the rules used to accomplish those goals. Therefore, a metadata schema should be analysed in terms of its intended objectives and the principles that guide its implementation.

Domains include environmental domains, object class domains, and object format domains. Environmental domains relate to the discipline or community served by the schema, object class domains refer to groups of similar objects, and object format domains focus on the composition or format of the object. These domains are important because they allow metadata categories to be defined and created in multiple ways.

Architectural layout focuses on the structure of a metadata schema, including its component parts, the number of metadata elements, and the level of granularity reflected in the refinement and specification of those elements. In this study, particular attention was given to object class domains and architectural layout because of their relevance to the research objectives. Metadata creation for oral history content should support accessibility by providing multiple access points and descriptive elements that facilitate the discovery and retrieval of information. The next section discusses the theory relating to access points for improving access to oral history content.

1.4.2 The Theory of Metadata Enriching and Filtering

The Theory of Metadata Enriching and Filtering suggests a mixed metadata approach, where metadata experts provide the basic descriptive metadata, structure, and interoperability (a priori metadata), and users continue adding their own interpretations (post-hoc metadata) (Alemu 2015). The theory emerged from the following four overarching metadata principles: enriching, linking, openness, and filtering (Alemu, Stevens, Ross & Chandler 2014). Figure 1.5 illustrates these four principles.

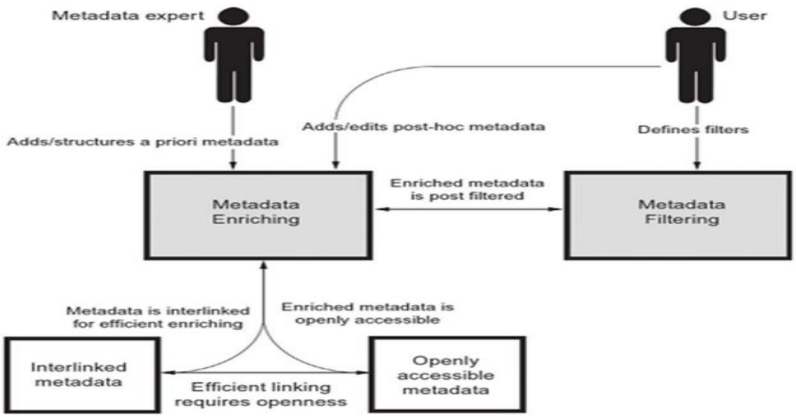


Figure 1.5: The Theory of Metadata Enriching & Filtering (Alemu 2015)

The theory of metadata enrichment and filtering places users at the centre of the metadata creation process and emphasises the role of metadata in improving the accessibility, discoverability, and retrieval of information resources (Alemu 2015). The theory proposes that users should not only access information objects through existing metadata but should also contribute to metadata creation by adding their own descriptions, interpretations, and terminology. According to Alemu (2015), enriching metadata with user-generated descriptions improves the representation of the aboutness of information objects. Gartner (2016) further argues that enriched metadata should be presented to users in a simple, meaningful, and user-friendly manner based on their information needs.

The theory is based on four main principles (Alemu, Stevens, Ross & Chandler 2014; Alemu 2015):

The Principle of metadata enrichment involves continuously enhancing expert-created metadata with user-generated metadata. This principle aligns with the concept of folksonomies, where users label or tag information objects using their own terms (Avery 2010). In archival environments, folksonomies can complement traditional classification systems and promote user participation in the organisation of information (Silva 2017).

The principle of linking refers to connecting metadata across different sources. This principle ensures that metadata are uniquely identified, structured, and interconnected, allowing information from multiple sources to be integrated. Metadata openness promotes the sharing

of metadata in open, reusable, and recombinable formats. Institutions are encouraged to make metadata freely available to support broader access and reuse.

The principle of filtering ensures that enriched metadata are presented according to individual user needs. To enhance the findability and usability of information objects, irrelevant search results and improving contextual relevance should be filtered. This suggests that enriched metadata should be linked, openly accessible, and presented in ways that meet users' specific needs (Alemu 2015). While metadata enrichment increases the amount of descriptive information available, filtering helps simplify its presentation through a user-centred approach.

This theory is particularly relevant to oral history content because metadata schemas should facilitate access, discovery, and retrieval. Accordingly, the principles of enrichment and filtering provided a useful framework for analysing the metadata schema used at the National Film, Video and Sound Archives (NFVSA). Furthermore, the standardisation of metadata practices in archives and other library and information services (LIS) institutions remains essential for effective metadata creation and application. The next section therefore discusses international metadata standards relevant to managing access to oral history content. The main purpose of metadata is to support the accessibility, discoverability, and retrievability of information resources.

Since this study contends that the metadata schema used for oral history content should facilitate access to this oral history content, the principles of enrichment and filtering appear to be applicable to the creation of metadata for oral history content. This framework assisted the researcher in analysing and describing the metadata schema used in NFVSA. Standardisation of practices relating to metadata creation in archives and other LIS centres is important for the metadata schema and application.

1.4.3 The Open Archives Information System framework

The Open Archival Information System (OAIS) is a reference model developed by the Consultative Committee for Space Data Systems (CCSDS) to guide the establishment and management of digital preservation repositories (Bettivia 2016). Standardised in 2002, the OAIS framework provides a common vocabulary and conceptual structure for understanding archival environments, their functions, and interactions (Tomášek 2018; Donaldson, Zegler-Poleska & Yarmey 2020). It describes the environment, information objects, and functional

components required for the long-term preservation of digital materials (CCSDS 2019; ISO 14721 2012):

a)OAIS Environment

The OAIS environment consists of four entities: Producers, Consumers, Management, and the Archive. Producers supply information to the archive, while Consumers access and use preserved information. The Designated Community refers to the group of Consumers expected to understand the preserved content. Management is responsible for setting policies, determining what information should be preserved, and securing resources for the archive. The archive itself carries out the day-to-day preservation activities (Lavoie 2000; ISO 14721 2012).

b)OAIS Information Objects

OAIS identifies four types of information objects: Content Information, Preservation Description Information (PDI), Packaging Information, and Descriptive Information. Content Information consists of the data object and the Representation Information needed to understand it. PDI supports preservation and includes Reference Information (identifiers such as ISBNs and URNs), Provenance Information (history and custody of the content), Context Information (relationships and purpose), and Fixity Information (checksums and digital signatures that verify integrity).

Packaging Information binds all components into a single identifiable information package, while Descriptive Information contains metadata and finding aids that support discovery and retrieval. OAIS also defines three information packages: the Submission Information Package (SIP) submitted by the Producer, the Archival Information Package (AIP) preserved within the archive, and the Dissemination Information Package (DIP) delivered to the Consumer. The DIP may be identical to or different from the AIP depending on access requirements.

c)OAIS Functional Entities

The OAIS model defines six functional entities:

Ingest receives SIPs, validates them, creates AIPs, and extracts descriptive metadata.

Archival Storage stores, maintains, and retrieves AIPs while ensuring their long-term preservation.

Data Management manages descriptive information and supports searching, reporting, and database updates.

Access enables users to search for information, retrieve relevant metadata, and obtain DIPs.
Preservation Planning monitors technological changes and develops strategies to address obsolescence through migration or emulation.
Administration manages daily operations, develops policies and standards, oversees agreements with producers, and ensures compliance with archival requirements.

Overall, the OAIS model provides a comprehensive framework for digital preservation and has been widely used to identify and manage preservation metadata (Lavoie 2000; Bettivia 2016; Tomášek 2018; Donaldson, Zegler-Poleska & Yarmey 2020). However, because preservation metadata was not the primary focus of this study, the model was considered only partially applicable.

1.4.4 FAIR Principles

The FAIR Principles were developed to improve the Findability, Accessibility, Interoperability, and Reusability of digital resources (Wilkinson et al. 2016). Unlike OAIS, which focuses on preservation, FAIR emphasises machine-actionability, enabling computer systems to locate, access, and reuse data with minimal human intervention (The FAIR Guiding Principles 2016). To be accessible, data and metadata should be available through open and implementable protocols, while metadata should remain accessible even when data are no longer available. To be interoperable, metadata should use shared vocabularies and standard languages and maintain links to related metadata. To be reusable, metadata should be richly described, include provenance information, provide clear usage licences, and comply with community standards (Wilkinson et al. 2016).

Metadata support retrieval, interoperability, and reuse through standardised communication processes. Although this study primarily focused on metadata access rather than interoperability and reuse, the FAIR Principles partially align to this study, as the focus will be on analysing metadata schema for metadata access and minimal attention to interoperability and reusability of the metadata. The researcher acknowledges that the principles cannot be separated and are interlinked, however, the scope of this study limits itself to describing and analysing the metadata schema used for metadata creation to access oral history content.

1.4.5 Open Archives Initiative Protocol for Metadata Harvesting

The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) was established in 2001 as a low-barrier standard for repository interoperability and metadata exchange over the web using XML and HTTP (Kapidakis 2018). Metadata harvesting involves the automated collection of metadata from multiple repositories to support discovery, sharing, and access across digital collections (Baca 2016; Corrado 2018). OAI-PMH supports interoperability through six services or "verbs". The OAI-PMH defines six 'verbs' as the components of communication between data sources and service providers (Openarchives 2015). The verbs include Identify, List Sets, List Metadata Format, List identifiers, Get Record, and List Records (Table 1.1), invoked within the HTTP.

Table 1.1 OAI-PMH Six verbs/services (Nelson 2001).

Verb	Description
Identify	returns a description of the archive
List Search List	returns a list of sets in use by the archive
List metadata formats	returns a list of metadata formats used by the archive
List Identifiers	returns a list of record ids (possibly matching some criteria) given a record id
Get Records	returns that record
List Records	returns a list of records (possibly matching some criteria)

The first three verbs provide information about the repository itself, while the latter three support metadata harvesting and retrieval. OAI-PMH enables repositories to exchange metadata and supports cross-repository discovery services (Corrado & Moulaison 2014; SAA 2023). Although OAI-PMH plays an important role in metadata interoperability and exchange, it was not considered directly applicable to this study because the research focused on the metadata schema used to provide access to oral history content, rather than on metadata harvesting or exchange between repositories.

1.4.6 International Standards Organisation

Standards provide benchmarks for measuring quality and guiding professional practice. They assist organisations in deciding whether to adopt, adapt, or develop metadata schemas to meet

their information management needs (SAA 2023). Several international standards are relevant to metadata creation, management, and access.

ISO 14721:2012 – Space Data and Information Transfer Systems (OAIS) provides a framework for the long-term preservation and access of digital information. It enables comparisons of archival architectures, preservation strategies, and data models, while promoting a shared understanding of digital preservation practices. The standard also supports the development of other OAIS-related preservation standards.

ISO 15489 – Information and Documentation: Records Management establishes principles for creating, capturing, managing, and organising records in any format or environment. The standard addresses records systems, metadata, policies, responsibilities, business requirements, and procedures for maintaining records throughout their lifecycle.

ISO 23081-1 – Information and Documentation: Managing Metadata for Records focuses on the standardised description of records and their contextual information. It aims to support interoperability, facilitate the exchange of records across systems, and encourage the consistent reuse of metadata over time and across applications (SANS 2010). The standard aligns with ISO/IEC 11179.

ISO/IEC 11179 – Metadata Registries provides guidance on standardising and registering data elements to ensure that they are understandable and interoperable across systems and organisations. The standard covers the framework, classification, registry models, data definitions, naming conventions, and registration procedures for metadata elements.

ISO 25964-1 – Information and Documentation: Thesauri and Interoperability with Other Vocabularies provides guidelines for developing and maintaining thesauri for information retrieval. It applies to all forms of information resources, including text, audio, video, images, and multimedia collections (ISO 25964 2013). The standard also promotes interoperability between thesauri and other knowledge organisation systems, such as taxonomies, ontologies, classification schemes, and authority lists (Raghavan & Krishnamurthy 2013).

These standards are relevant to exploring metadata schemas for oral history content at the NFVSA because they provide guidance for enhancing metadata quality, usability,

interoperability, and access. They also support decisions on whether existing metadata schemas should be adapted or new frameworks developed.

ISO 9000 Family – Quality Management Standards focuses on organisational quality and continual improvement. The standards emphasise customer focus, leadership, engagement of people, process approaches, improvement, evidence-based decision-making, and relationship management. In particular, ISO 9001 promotes effective processes and skilled personnel to improve organisational performance and meet user expectations. For the NFVSA, the implementation of quality management principles is important for ensuring effective metadata creation, knowledgeable staff, and improved access to oral history content. Ultimately, the quality of metadata and staff expertise contributes significantly to user satisfaction and the successful fulfilment of the institution's mandate.

1.4.7 Policies and guidelines in archives

A policy framework is essential for guiding organisational functions and ensuring effective management practices (Stephanie & Ou 2022). Policies are particularly important for metadata management and access to oral history records. Although legislation exists, archival legislation in many ESARBICA countries remains outdated (Msibi 2015). South Africa has a strong legislative framework for archives and records management; however, as Katuu and Van der Waldt (2016) note, there are often no specific and clear guidelines for specialised records, a challenge that also applies to metadata for oral history content.

The primary legislation is the National Archives and Records Service of South Africa (NARSSA) Act of 1996, which mandates the collection, preservation, and dissemination of public records. It is supported by the Promotion of Access to Information Act (PAIA) of 2000, which guarantees public access to information held by public and private bodies, and the Electronic Communications and Transactions Act (ECTA) of 2002, which gives legal recognition to electronic records (Katuu & Van der Waldt 2016). Additional guidance is provided through SANS 15489 and the Records Management Policy Framework for the Public Service.

Despite this legislative framework, challenges remain. Goh (2016) found that archival legislation often lags behind technological developments, while Nsibirwa (2019) argues that legislation and archival practices in developing countries have not kept pace with technological change. Netshakhuma (2019) identified gaps in records management legislation relating to records transfer, disposal, and accessibility. Similarly, Tembe and Nsibirwa (2022) note that existing legislation provides limited guidance on the acquisition, management, preservation, and access of oral history records.

The NFVSA operates within legislation such as the NARSSA Act, PAIA, POPIA (2013), and the Promotion of Administrative Justice Act (PAJA) of 2000, as well as the NARSSA Managing Electronic Records Metadata Requirements and NARSSA Regulations (2002). These frameworks support lawful, accountable, and transparent management of records, including metadata creation and access.

Overall, South Africa has a robust legislative framework for archives and records management. However, there is limited policy guidance specifically addressing metadata and access to oral history content, highlighting the need to examine metadata schemas and related policies at the NFVSA.

1.4.8 Oral History Association and oral history best practises

The Oral History Association (OHA) is an international organisation that promotes best practices in oral history and supports oral historians through engagement with policymakers, educators, and other stakeholders. It encourages communication among members and establishes standards for the collection, preservation, dissemination, and use of oral history content. To support these activities, the OHA has developed objectives, guidelines, and evaluation standards for oral history practice.

The OHA Best Practices Manual (2019) provides guiding principles for those involved in the collection, interpretation, preservation, and use of oral histories. Recognising the diverse nature of oral history practice, the OHA emphasises that oral history interviews are primary sources that should be accessible, discoverable, and understandable. Consequently, oral history collections require specific archival principles and best practices to support their management and use.

The Principles and Best Practices Manual (OHA 2019) is not intended as a one-size-fits-all guide. Archivists and oral history practitioners are encouraged to apply the practices most appropriate to their organisational or community contexts. Metadata and description are particularly important because they provide contextual information that enhances the discovery and accessibility of oral history content. The Oral History Association Metadata Task Force (OHA MTF) has developed detailed guidelines and tools to support metadata creation and description. These guidelines offer a foundation for developing metadata and descriptive practices tailored to oral history collections.

Metadata and related concepts were developed and aligned with the review of the literature and theories that underlie this study. These concepts will be fully explained in chapter two as part of the review of the literature. The conceptual framework of this study guided this study, which also guided the boundaries and originality of this study. Therefore, the next section focuses on the originality of the study.

1.5 ORIGINALITY OF THE STUDY

All institutional regulations require that research at doctoral level be original and significant. The originality of a study is determined by how much it adds to knowledge – that is, by how much the researcher's work sheds light on, and deepens our understanding of, a certain topic (Baptista, Frick, Holley, Remmik, Tesch & Âkerlind 2015). In addition, Maria, Cruz, San Juan, Tan and Amarnani (2022) assert that originality in human sciences suggests that the study should contribute to existing knowledge, either by applying a different perspective, or by exploring a new area of study; recognised by others as original; and that a new coherence has to be developed in the work which is not found in any existing work. Borrowed from Guetzkow, Lamont and Mallard (2004), seven types of originality are recognised for use in social and human sciences research. These include the original approach, the understudied area, the original topic, the original theory, the original method, the original data, and the original results (Guetzkow et al. 2004). Similarly, Francis in Phillips and Pugh (2005) argue that originality could be achieved by establishing a piece of new information or by testing someone’s idea. The originality of this study is thus based on the setting of a major piece of new information and the testing of someone’s ideas.

Little research has been reported in the literature on metadata schemas and the evaluation of such metadata in the LIS sector. From the reviewed literature, this is the first study to investigate the metadata schema in South African archives, which lies in the originality of this study. The researcher is of the opinion that this study could emphasise the framework of a metadata schema for oral history content. The research is also aware of the NARSSA metadata principle, the EAD principles, and the OHA manual(2019) for the description of oral history content. This study makes a decent contribution to the literature on this important aspect of research and provides useful and practical information in the field of archives. Therefore, the study intends to contribute to the establishment of a new framework for the development process of metadata schema in South Africa that practitioners can use to create their metadata records.

This study focused on metadata access, using oral history content as a case to determine how metadata can be better applied to provide more efficient access. Archivists should be able to take the possible solutions and/or framework from this study and apply them not just to oral history content in their own collections, but also to other archival collections, to improve and/or enhance discoverability and accessibility thereof.

1.6 JUSTIFICATION OF THE STUDY

According to Babbie (2013), the importance of the study offers the researcher an opportunity to discuss the contribution of the study. This study identifies and suggests solutions to the metadata classification schema and related problems. With identified problems at their disposal, information centres in general, and archival centres in particular, will be able to improve metadata practices, which would lead to improved access to archival materials. This study also educates South African archives in general about the importance of metadata classification and how it can be used to improve access to oral history content and, subsequently, improve metadata creation, thus increasing the number of users of oral history content. The study will help any other information centres to have a better understanding of the organisation and creation of sources and the retrieval of information in their organisations. The study is also likely to form the basis for further research on the metadata schema of oral history content in LIS centres in South Africa. Therefore, the study has the potential to be additional work in the LIS field, as the intention is to publish certain sections of the study in an accredited journal.

This research is significant in understanding the diverse ways in which archivists, librarians, and oral historians provide descriptive content for the oral history collections for which they are responsible. This includes descriptions and metadata for analogue and digital oral history collections. This research can also reveal descriptions and metadata schemas, which fields or elements each collection uses, and the ways in which the various fields or elements are similar and different across collections. The author hopes that exploring the method and content of describing oral history collections will help establish useful ways to promote access to, and increase the use of, these materials. With this knowledge, those working with oral history collections could maximise the value and benefits of collections for librarians, archivists, and citizens using these collections. This study aimed to contribute to the proposal of a metadata framework to be adopted in the description of oral history collections nationally and more broadly. Therefore, to channel the focus and importance of this study, it is relevant to outline the scope and delimitation of this study.

1.7 SCOPE AND DELIMITATION OF THE STUDY

Archives in South Africa are usually classified into two categories. One category is government and/or public archives, responsible for official records of the state, provinces, municipalities, oral history content, and others. The second category is academic, private, or community archives for manuscript collections, primarily of the founding (Swanson 2015). The following are some centres where oral history content is housed:

- Robben Island Museum: The Robben Island Museum houses an oral history collection that documents the experiences of political prisoners who were held on the island during apartheid. The collection includes more than 1,500 recordings, some of which have been made available online.
- University of Cape Town: The University of Cape Town has the Centre for Popular Memory that collects and digitises oral history interviews with people who have lived through significant historical events in South Africa. The centre has made some of its collections available online.
- South African History Archive: The South African History Archive is an independent archive that collects and preserves historical documents, including oral history interviews related to the anti-apartheid struggle. This archive has made some of its collections available online.

- Community-based archives: There are also several community archives in South Africa that collect and preserve oral history interviews related to specific communities or regions. For example, the KwaZulu-Natal Oral History Project collects and preserves oral histories of people from the province of KwaZulu-Natal, and the District Six Museum collects and preserves oral histories of people who were forcibly removed from their homes in the District Six neighbourhood of Cape Town during apartheid.

In addition to these institutions, there are also several initiatives aimed at promoting access to oral history collections in South Africa. For example, OHASA promotes the collection, preservation, and use of oral history materials and provides training and support to oral historians and archivists. This study focused on the public archives in South Africa, with a specific focus on the NFVSA as custodians of these oral history collections.

While the NFVSA is responsible for legal deposit, film, audio-visual, and sound archives, this study only focused on oral history acquired at this institution in whatever format. While the researcher acknowledged that there are various factors that might affect access to oral history oral conten, this study confined itself to metadata creation and metadata schema boundaries. Archivists must analyse, classify, describe, and organise the materials in their collections. Various factors This study also focused on archival professionals who are tasked with metadata in the NFVSA and those tasked with oral history content. The researcher was also aware that the metadata records include various formats of oral history content. Therefore, the purpose of the study was to investigate metadata, which is the basic tool for gaining access to oral history content. Since the purpose of this research was to assist in developing a framework to ensure oral history content and its retrievability, oral history metadata records was best positioned to assist in achieving the main purpose of the research.

This study took into account the fact that metadata records are created for archival materials, in general. It is correct to view the metadata schema as used for both preservation and access. However, this study focused on metadata to access oral history content.

1.8 LITERATURE REVIEW

Bertram and Christiansen (2017:13) define a literature review as “a comprehensive overview of previous research studies on a subject, providing background knowledge on studies that have

been examined before." To identify and analyse existing scholarship, the researcher consulted books, journals, theses, dissertations, online databases such as EBSCOhost and Taylor & Francis, and scholarly sources available through Google Scholar. A literature review also serves to identify theories and models relevant to the research topic (Du Plooy-Cilliers, Davis & Bezuidenhout 2014:101). Creswell and Creswell (2018) further emphasise that the identification of appropriate theoretical and conceptual frameworks strengthens the clarity and effectiveness of research.

Metadata schemas are essential for the discovery, description, and accessibility of information resources (Sulehri & Warraich 2020). Metadata provide the foundation for effective resource discovery, access, and sharing across distributed digital collections (Park & Tosaka 2010:104). They are structured data that support functions such as preservation, analysis, reformatting, discovery, and use of resources (IASA 2021). The literature reviewed in this study focused on metadata schemas, metadata standards, access to oral history content, and technologies that support metadata creation and access. These themes are discussed in detail in Chapter Two. The review enabled the researcher to identify gaps in the fields of oral history and archives and to determine research methodologies appropriate to the study. The next section outlines the research methodology adopted for this research.

1.9 RESEARCH METHODOLOGY

The research methodology outlines the procedures followed when conducting research. Brynard, Hanekom and Brynard (2014:38) state that "research methodology focuses on the process of research and the decisions that the researcher has to take to execute the research project." According to Wagner, Kawulich and Garner (2012:52), a research methodology is "a section where assumptions about the nature of reality and knowledge, values and theory and practise on a given topic come together." Research methodology includes aspects such as research design, research approach, population, sampling, data collection, data analysis, and quality criteria for this study.This study adopted inductive reasoning within the constructivism/interpretivism philosophical framework of the qualitative approach. Data were collected from document analysis and interviews with NFVSA staff and analysed through thematic analysis. To ensure the quality of this study, the researcher addressed credibility, transferability, confirmability, and dependability within the qualitative approach. Reliability, validity, and objectivity were addressed in this study, despite them being commonly used for

quantitative studies, but also recognised as a need in qualitative research. A comprehensive description of the research methodology is discussed in chapter three of this thesis.

For each study conducted, there are findings and recommendations to address the problem at hand. Because research is conducted to provide plans and strategies to address specific problems, as was the case in this study, the result of such research should be distributed and made known. Therefore, it is valuable to explain the dissemination of the findings of this study.

1.10 DISSEMINATION OF THE FINDINGS

Ocholla (1999:141) argues that "the possession of information without it being well disseminated is useless and research is incomplete until it is disseminated." In line with this view, the findings of this study will be disseminated through presentations at local, regional, and international conferences, workshops, and seminars. The completed dissertation will be made available through the UNISA Institutional Repository (IR), while an electronic copy will be shared with the National Film, Video and Sound Archives (NFVSA).

To ensure broader access to the research outcomes, copies of the study will be submitted to the UNISA Library, the participating archival institution, and made available to interested participants. In addition, journal articles derived from the dissertation will be published in accredited national and international scholarly journals.

Creswell (2014) emphasises that researchers remain accountable for the information they disseminate, regardless of the outcomes of the study. Consequently, the dissemination of scientifically credible and authoritative findings is essential to ensure that the research benefits its intended audience. Given the importance of responsible research practice, ethical considerations are discussed in the following section.

1. 11 ETHICAL CONSIDERATIONS

Researchers are expected to identify and address ethical considerations in every study. Research ethics are essential for protecting the dignity of participants and safeguarding the information generated through research (Fouka & Mantzorou 2011). According to Gajjar (2013), adherence to ethical principles promotes research objectives, advances knowledge, minimises errors, supports collaboration among researchers and institutions, and ensures

accountability to the public. Polit and Beck (2014) further emphasise that the credibility of research, public trust, and acceptance of research findings depend on adherence to ethical standards and the authenticity of the information reported. In this study, ethical compliance was ensured through the acquisition of ethical clearance from the University of South Africa (Appendix C) before data collection commenced.

Arifin (2018:30) argues that "considering ethical issues is crucial throughout all stages of qualitative study to maintain the balance between the potential risks of research and the likely benefits of the research." This is particularly important in qualitative research, where researchers are required to observe, interpret, and make sense of participants' experiences and perspectives (Sanjari, Bahramnezhad, Fomani, Shoghi & Cheraghi 2014). Key ethical principles in qualitative research include informed consent, anonymity, confidentiality, privacy, and scientific integrity (Wiles 2013). As this study adopted a qualitative approach, these ethical considerations were observed throughout the research process and are discussed in detail in Chapter Three.

For the purpose of this study, it is correct to discuss the different terms that were used throughout the study and how they were be used. Therefore, the next section focuses on the definition of concepts.

1.12 DEFINITION OF CONCEPTS

In this section, the researcher defines and explains the main terms used in this study.

1.12.1 Metadata

Metadata are commonly defined as data about data and refer to structured information that describes, identifies, and provides context for information resources (Rawson 2018; Assmann 2009). Gilliland (2000) defines metadata as "the sum total of what one can say about any information object at any level of aggregation," while Gilliland-Swetland (2000) notes that metadata identify resources, describe their content and creators, and establish relationships with other resources. Similarly, Grassi et al. (2018) define metadata as structured information that describes, explains, locates, manages, preserves, and facilitates the retrieval and use of information resources over time. For this study, metadata are understood as structured

information that describes and supports access to information resources regardless of their formats.

1.12.2 Metadata schema

Metadata schema is defined as a logical plan showing the relationships between metadata elements, normally through establishing rules for the use and management of metadata, specifically with regard the semantics, the syntax, and the optionality (obligation level) of values (ISO 23081). This suggests that the metadata schema refers to the overall structure of the metadata, as it describes how the metadata are set up and usually addresses standards for common elements to be described. Metadata schemas provide standardised methods to describe the content of oral history recordings, such as the names of interviewees, dates, locations, and subjects.

11.2.3 Metadata standard

The metadata standard includes principles and implementation issues to tag information and ensure consistency in metadata application. This means that metadata standards establish the rules for using those elements to describe a resource. Zhang and Gourley (2009) define "metadata standards as specifications or definitions that establish common meaning, usage, and interpretation of metadata and are applied to electronically stored data and used to describe the definition, structure, and administration of data files with all contents in context for easy use of the captured and retrieved data for further use (Xie & Matusiak 2016). Although the metadata standard and metadata schema are often used interchangeably, in this study, the two concepts are used independently, but as related to each other.

1.12.4 Oral history content

Oral history involves recording and preserving personal experiences, memories, and information through interviews and narratives (Goodchild, Ambrose & Maye-Banbury 2017; Thomson 2011). Unlike storytelling, which focuses on sharing stories and experiences from personal perspectives (Chancellor & Lee 2016; Albarelli & Starecheski 2013), oral history is a systematic process of documenting lived experiences for historical purposes.

Oral history refers both to the process of conducting recorded interviews and to the records produced from those interviews (Pryse et al. 2021). Butler (2008:34) defines it as "a recorded interview of an individual or group of individuals by a historian, researcher, or other interested individual doing the interview." Oral history content serves as a primary source of historical and cultural evidence and should be preserved, accessible, discoverable, and understandable for current and future use (OHA 2019).

1.12.5 Access to oral history

Simply put, 'access' refers to the freedom or ability to obtain or use something (Merriam Webster Dictionary). Therefore, access to oral history may refer to the ability and availability of oral history to be used. According to OHA (2019), oral history access allows the public to interact with oral history content. Permissions and ethical use, transcription, and public access and discoverability are some of the key components considered in the access and use of oral history (OHA 2019). On the other hand, oral history access should include retrievability through search aids.

1.13 STRUCTURE OF THE THESIS

This study is structured as follows:

Chapter one: Setting the scene

This chapter provides the background of the study, the research problem, the research purpose, the justification, and the originality of the study. It also provides contextualised definitions of key terms.

Chapter two: Literature review

This chapter provides an overview of the literature reviewed for studying classification schemas and oral histories. The review of the related literature placed the study in context and highlighted important trends in metadata creation, standards in general, and classification.

Chapter three: Research methodology

This chapter explains the research methodology used to study the metadata schema for oral history content. The discussion covers the following: research paradigm, research design,

research approach, population, and sample, as well as data collection and analysis. This chapter also explains the research ethics to which the researcher adhered.

Chapter four: Data presentation and analysis

This chapter outlines the presentation and analysis of the data. Data collected from document analysis and interviews are analysed.

Chapter five: Interpretation and discussion of data

The results of the study are discussed and linked to the existing literature and the objectives of the study. This chapter focuses on discussions and interpretation of the results presented in chapter four.

Chapter six: Summary, conclusions, and recommendations

This chapter summarises the key findings of the study and outlines the implications of this study for research and practice. In addition, the chapter makes recommendations based on the findings. Finally, conclusions were drawn from the study findings.

1.14 SUMMARY

This chapter discussed the following aspects of the study: background, research problem, research purpose, conceptual framework, originality of the study, justification of the study, scope and delimitations, outline of the literature review, research methodology, dissemination of the findings, ethical considerations, as well as definition of concepts. Furthermore, the chapter outlines how the chapters are organised in this study. The next chapter discusses the literature review on metadata schema and access to oral history content.

CHAPTER TWO
LITERATURE REVIEW

2.1 INTRODUCTION

Chapter One provided the foundation for the study by outlining the background, problem statement, purpose, conceptual framework, significance, scope, research methodology, ethical considerations, key concepts, and structure of the thesis. This chapter reviews the literature on metadata schemas and access to oral history content in archival institutions. A literature review is an essential component of research because it situates a study within existing knowledge, identifies relevant theories and models, and highlights gaps that require further investigation (Ngulube 2018; Creswell & Creswell 2018). It also assists researchers in understanding previous work on a topic and in establishing the theoretical and conceptual foundations of a study (Flick 2018).

According to Du Plooy-Cilliers, Davis and Bezuidenhout (2014), a literature review involves the systematic identification, analysis, evaluation, and synthesis of relevant scholarly sources. Similarly, Saunders, Lewis and Thornhill (2019) regard it as a critical process that supports the achievement of research objectives by providing a foundation for understanding and interpreting a research problem. In qualitative research, the literature is used to frame the study, inform the conceptual framework, and compare findings with existing knowledge (Creswell 2014).

The review of literature in this chapter focuses on metadata schemas, metadata standards, oral history content, access to oral history collections, and related technologies. It also examines relevant theories, models, standards, and policies that inform metadata creation and access. Through this review, the study identifies gaps in the existing literature and establishes the context for investigating metadata schemas for oral history content at the National Film, Video and Sound Archives (NFVSA). The following section presents the literature map for the study.

2.2 LITERATURE MAP FOR THIS STUDY

A **literature map** organises and categorises existing literature according to a broad framework, such as theory, theme, or purpose (Aurini, Heath & Howells 2016:24). In this study, the literature was reviewed to examine metadata schemas for oral history content and to understand

the relationship between metadata schemas, metadata standards, and access to oral history resources.

The literature search was conducted using databases such as **EBSCOhost** and scholarly sources including **Google Scholar**. Search terms included *metadata*, *metadata standards*, *oral history content*, *metadata schema and description of audio-visual resources*, *metadata schema and standards*, and *access to oral history*. According to Creswell (2014), a literature review requires the identification, selection, critical evaluation, and synthesis of relevant sources to establish how a study contributes to existing knowledge. To achieve this, the literature must be systematically organised through a literature map.

The literature map (Figure 2.1) provides a visual overview of the key concepts and relationships underpinning the study of metadata schemas and access to oral history content (Creswell 2014).

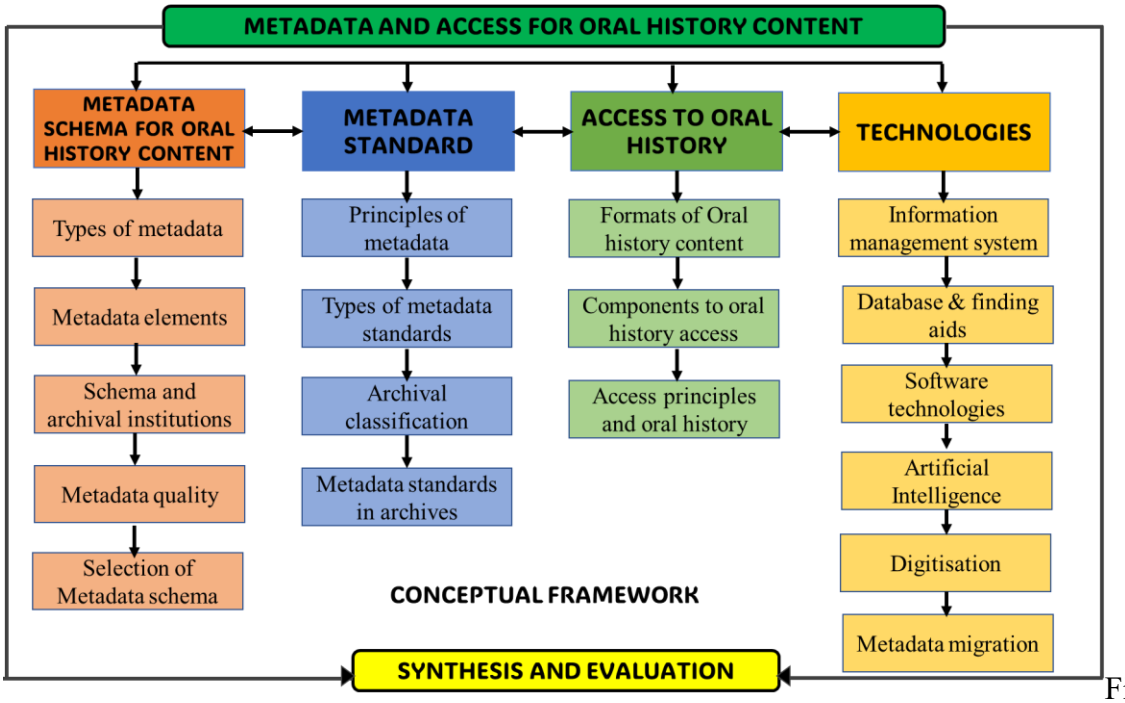


Figure **2.1: Literature Map (Researcher's own work)**

It is necessary to contextualise this study, so the literature is organised around the research objectives, which are subdivided into subheadings. Both metadata schema and metadata standards are important in organising and describing metadata; however, they refer to different aspects of metadata. A metadata schema defines the structure and content of metadata, while a metadata standard specifies how those metadata should be created, managed, and used. As a

result, the purpose of this study was to explore the metadata schema used for oral history content. To put this study into perspective, it is necessary to begin this chapter by putting oral history content into context.

2.3 ORAL HISTORY CONTENT IN ARCHIVAL INSTITUTIONS

Oral history content is a valuable source of cultural heritage and evidence of both historical and everyday experiences (Tebeau 2013). Traditionally, oral histories were preserved in archives and historical organisations using analogue formats such as audio cassettes. However, digital technologies have created new opportunities for the organisation, preservation, and dissemination of these resources (Xie & Matusiak 2016). Boyd (2011) notes that digital recording technologies and digital collections have expanded the ways in which life stories can be captured and shared with wider audiences. Digitisation also supports preservation by reducing the risks associated with fragile and deteriorating analogue recordings (Xie & Matusiak 2016).

Despite these advantages, oral history content remains vulnerable to loss. Xie and Matusiak (2016) argue that oral history collections are continually threatened by human error, technological failure, natural disasters, and other external factors. Similarly, Murillo (2014) identified deterioration, format obsolescence, and insufficient metadata for discovery as major threats to the preservation of digital information. Patterton, Bothma and Van Deventer (2024) define data at risk as data that may be lost because of media deterioration, inadequate documentation, or storage in obsolete or non-digital formats.

According to Patterton (2023), factors contributing to data at risk include inadequate backups, poor storage, lack of continuity and data management plans, obsolete technologies, equipment failure, incompatible software, and insufficient metadata. Among these, the lack of metadata is particularly significant because it limits the discovery, accessibility, and use of information resources.

Oral history collections preserved in archives are intended to support access by researchers, learners, and communities and may be used for historical, educational, cultural, and creative purposes. While recognising the various risks that threaten oral history content, this study focuses specifically on metadata and its role in facilitating access. Given that inadequate

metadata can hinder the discovery and retrieval of oral history resources, the study examines metadata schemas for oral history content. The next section therefore discusses metadata schemas for oral history collections.

2.4 METADATA SCHEMA FOR ORAL HISTORY CONTENT

The objective of this study was to describe the metadata schema used for oral history content. The conceptual framework of this study was to confirm the metadata structure that fit the metadata creation for oral history content. Metadata are fundamental to the discovery and access to special collections (Sandaver 2020) such as oral history content. Schemas are an important part of metadata. The metadata schema describes the overall structure of the metadata (Baca 2016). The type of metadata schema used is determined by the available metadata (Pearce-Moses 2005; Berry 2018). As a result, it is necessary to first discuss the types of metadata. The following section provides a brief description of each type of metadata.

2.4.1 Types of metadata

Metadata provides a consistent and logical means to maintain records accessible and useable across time. As such, it has to be broadly acknowledged, adopted, and curators of oral history information should be encouraged to utilise it (Groenewald & Breytenbach 2011). Accurate metadata is essential for a search function, particularly when functionality is the primary objective, as without correct metadata content cannot be readily retrieved (De la Porte & Higgs 2019). To ensure long-term accessibility, all discoverability and access functions rely on the available metadata, or information that describes collections in the repository (Berry 2018). The type of metadata also influences its availability. Metadata are classified into three types, including but not limited to, administrative, descriptive, and structural metadata (Berry 2018; OHA 2019). These three types of metadata are critical for archival material access and retrieval (Ramesh, Vivekavardhan & Bharathi 2015; ICA 2016; OHASA 2019), so they were the focus of this study. The three types of metadata that oral historians should collect are summarised in the following section.

2.4.1.1 Administrative metadata

Administrative metadata include information aimed to facilitate resource management, such as the date of creation, the rights and restrictions of access, and archiving, control or processing

activities (Riley 2017). Administrative metadata elements include (Ramesh, Vivekavardhan & Bharathi 2015):

- Preservation metadata to represent the recording of events
- Related to curation, submission, and dissemination and event history data, for linking with digital objects
- Metadata for rights management
- Technical information such as storage and format data
- Details of representation information including internal coding and rendering data

Given the nature of oral history content, it is significant that administrative metadata are included for accessibility. Administrative metadata are significant when collecting oral history content. It includes details on the rights and permission specified, who looked after the digital object, and the preservation methods that were employed. This comprises technical data that is relevant to all file types, such as digital signatures and checksums to ensure authenticity or specifics regarding the hardware and software that the digital object can be executed on (Weig & Slone 2018). Weig and Slone further indicate that content-type-specific information, like the width of an image or the duration of an audio file, are also technical details provided. Administrative metadata are also essential part of both digital and analogue oral histories. In the case of born digital oral history, these metadata are often generated by the capturing or recording device (Weig & Slone 2018).

2.5.1.2 Descriptive metadata

The main objective behind the creation of descriptive metadata records and context is to improve information access (Eíto-Brun 2015:72). Descriptive metadata facilitates the delivery and finding of digital content by providing features like author and title that cover the intellectual property. Additionally, it could offer historical context by identifying the print-based content that served as the initial source of a digital adaptation, for instance (Formenton & De Souza Gracioso 2022). Descriptive metadata are the primary source for users to access oral history content, and as such curatorial details such as associated dates, the name of the file, creator, description, summary, and locations, are provided (ICA 2016). Ramesh, Vivekavardhan and Bharathi (2015) add that descriptive metadata provide content data such as titles and subject terms; and context-related information related to the context and/or

creation; and useful in its identification and discovery. Ramesh, Vivekavardhan and Bharathi (2015) identify two descriptive metadata models, namely:

- Separate metadata model: Metadata are kept in a database linked to a resource by means of a URL or other descriptive metadata locator. Each metadata entry is an HTML page that is linked to the resource it describes.
- Embedded metadata model: Using metadata tags in the HTML page, the metadata creator embeds metadata into the web page as part of the content it contains.

Descriptive metadata are aimed at supporting and assisting users to discover, retrieve, and access oral history content. One may then conclude that descriptive metadata play an important role on how to find oral history content, how to distinguish from other similar collections, and bringing together related content, obtaining access, and evaluating, linking, and using oral history content (ICA 2016). This suggests that descriptive metadata remain the most relevant for access to oral history content as it allows to identify, locate, analyse, and administer images which form part of oral history content. Gilliland (2000:11-12) also refers to descriptive metadata as "a tool to describe, identify, and authenticate collections and oral history content in particular." In other words, descriptive metadata makes it easier for users to locate and understand oral history content collection (Riley 2017). However, as it involves more actions on interpretation and translation, the structural metadata remains crucial area of the metadata production.

2.5.1.3 Structural metadata

Structural metadata are information that describes how the many parts of complex information resources fit together, such as the text, image, and sound files that make up a multimedia web page or chapters in a book (ICA 2016). Gartner (2016) indicates that structural metadata capture both logical and physical linkages, such which image is embedded in which web page. One may thus refer to structural metadata as the recordings of the relationships that bind together digital items together.

The National Information Standards Organization (NISO) (Riley 2017) states that structural metadata describe the intellectual and/or physical components of a digital object, as well as its prospective layout. In a multifile digital object (for example, a scanned book with many page images), structural metadata describe the object's components and are usually linked to

packaging formats, like how oral history content descriptions are put together to form one metadata record.

Putting metadata into descriptive, structural, and administrative categories may be confusing, as such metadata elements are helpful to clarify the existing boundaries in any metadata schema. It is therefore necessary that the next section should focus on metadata elements.

2.4.2 Metadata elements

The metadata elements contain useful details about various situations. The conceptual framework of this study was designed to describe the metadata schema by determining elements that are applicable in the creation of metadata of oral history content. According to Yang (2016), the metadata element corresponds to a formally defined term to describe one of the properties (or attributes) of the resource of a certain type or with a particular purpose, such as the 'format' of a file. Grassi et al. (2018) propose a set of elements to support the semantics and syntax required to support interoperability. The metadata elements are subdivided into two components, namely attribute schema metadata and attribute value metadata (Grassi et al. 2018).

(a) Attribute schema metadata

Attribute schema metadata (ASM) refer to metadata for the attribute itself, that is, describing the format in which the attribute will be transmitted. ASM provides information that is applicable to the attribute being asserted, regardless of the value of that attribute. ASMs are intended to be static, discussed, and agreed upon in advance of the actual assertion. This will ensure that all materials are described with the same semantic understanding, allowing trust and interoperability of the represented attributes (Grassi et al. 2018).

(b) Attribute value metadata

Attribute value metadata (AVM) are elements that focus on the asserted value for the attribute. This element proposes values for metadata fields and a rationale for their inclusion. AVM can enable more informed decisions and support richer metadata creation.

The study by Hu, Ng and Xia (2018) revealed that users have raised the issue of interoperability among metadata schema as a way to improve the user experience of digital platforms that house cultural heritage information. These findings highlight the significance of specific metadata

elements as a perceived means of standardising oral history content . The next section focuses on the metadata schema.

2.4.3 Metadata schema in archival institutions

Metadata schemas are sets of metadata elements developed for a metadata creation purpose. Each scheme consists of a finite number of pieces, each of which specifies the content or syntactic rules (Formenton & De Souza Gracioso 2022). Jana and Rout (2022: 346) refers to metadata schema as a set of standardised structured metadata. Metadata schema describes how the metadata are set up and usually addresses standards for common components of metadata elements (ISO 2016). According to ISO (2016), a metadata schema is a framework or structure for organising and describing metadata in such a way that it defines the elements or fields that should be included in metadata and specifies how those elements should be used and formatted.

Although there are quite a number of schemas for metadata creation, for the purpose of this study, only the most popularly and widely used schema is discussed. Dublin Core Metadata Initiative (DCMI), Metadata Encoding and Transmission Standard (METS), Metadata Object Description Schema (Park & tosaka 2010; MacKay 2016; Jana & Rout 2022). Even though the MARC was the most widely used schema, it was excluded in this study as it is mostly used in libraries, and this study resides in an archival institution. Oral History Metadata Synchroniser (OHMS) was included in the study as is a widely used metadata schema for oral history collections (Wynne 2009). Gilliland (2000) refers to metadata schema as data structure standards which serves as categories and/or containers of data making up a metadata record. Access to Memory (AtoM) , even though regraded as a repository, will be discussed under this section as a structure standards providing elements or containers into which information is organised (SAA dictionary). The sections begins the discussion with DCMI, followed by METS, MODS, OHMS, as well as AtoM.

This section explains the metadata schema used in archival institutions.

a) Dublin Core

The Dublin Core is a set of metadata elements that provides a small and fundamental group of text elements through which most resources can be described and catalogued. Using only 15 base text fields, a Dublin Core metadata record can describe physical resources such as books,

digital materials such as video, sound, image, or text files; and composite media such as web pages (Ronzino, Amico & Niccolucci 2011). Dublin Core is an internationally used schema developed with the internet in mind; it is meant to describe digital items.

While not specifically designed for oral history collections, Dublin Core is a widely used schema for describing a range of digital resources, including oral history interviews. Dublin Core includes a set of basic metadata elements, such as title, creator, and date, which can be extended as needed to provide additional information about the oral history interview.

(b) Metadata Encoding and Transmission Standard

The Metadata Encoding and Transmission Standard (METS) is an XML schema developed expressly to serve as a general framework for storing all metadata related to a digital object (Corrado & Moulaison 2014). A METS file consists of four main sections: a file inventory that lists all of the files; an administrative metadata section that includes provenance and rights management information; a descriptive metadata section that includes bibliographic information; and a structural map that shows the user how the item's various components relate

(c) Metadata Object Description Schema and Transmission Standard

The Metadata Object Description Schema (MODS) is maintained by the Library of Congress and was developed to describe bibliographic concepts and metadata.(Rayan, Shimizu, Sieverding & Hitzler 2023) According to Zeng and Qin (2020: 414), MODS is "currently the most comprehensive descriptive metadata schema and which has already been mapped with Bibframe in order to enable reuse of MODS records in a BIBframe environment." Bibframe as a data model is flexible, adaptable, and allows easy integration of bibliographic data with other datasets and presents it in various formats, including linking to web resources. This schema may be used with any types of materials with any type of institution. Metadata schema was not designed with oral history content in mind, but it does allow metadata and data related to the object to be stored (and archived).

(d) The Oral History Metadata Synchroniser

The Oral History Metadata Synchroniser OHMS was designed to provide a standardised way of describing and accessing oral history content (Wynne 2009) and includes a range of metadata elements, including interviewee and interviewer information, date and location of the

interview, and subject headings. The OHMS also provides a time-stamped transcript of the interview, which can be linked to the audio or video recording, allowing users to search for specific topics within the interview. The OHMS also includes a set of guidelines for ensuring that metadata are consistent linked and complete to one another so that they can navigate the file.

(e) Access to Memory

Access to Memory (AtoM) allows standard-based archival description access in a multilingual, multi-repository setting (Noor 2020), enabling archival repositories to distribute their collections online easily and affordably. According to Kari and Barro (2016), AtoM generates metadata, retrieves diverse types of digital content, and may be tailored to meet the demands of different communities. Additionally, its collections are searchable, retrievable, and interoperable via the Web (Simons 2014). When a user flips the default template to a different description standard in Admin > Settings > Default templates, AtoM enables data to appear correctly in the new template whenever the selected standard has an analogous field (Noor 2020).

In general, there are several metadata schemas that can be used to describe oral history collections, each with its own strengths and limitations. The choice of schema will depend on factors such as the size and scope of the collection, the level of detail required, and the technical infrastructure available for storing and accessing the metadata. In short, the quality of the metadata is crucial. Therefore, it is relevant for the next section to focus on metadata quality.

2.4.4 Metadata quality and metadata creation

Metadata are a key element in the access structure, and as such, metadata quality is an important factor to ensure accessibility. If metadata quality is poor, the discoverability and accessibility will be compromised. Various authors agree on some of the characteristics of poor-quality metadata (Tani, Candela & Castelli 2013), which are categorised as follows:

(i) *Incorrect values:* These include inaccurate values; even when elements are applied successfully, metadata records contain values that do not accurately describe a particular resource.

(ii) *Incorrect elements*: These include inaccurate values; even when elements are applied successfully, metadata records contain values that do not accurately describe a particular resource.

(iii) *Missing information*: The record of metadata is incomplete.

(iv) *Information loss*: A few of the details' distinguishing characteristics are missing (this generally occurs during the process of converting metadata from one scheme to another).

(v) *Inconsistent values:* Distinct values linked to an element might accurately reflect a feature of the resource; yet, their discrepancies in written form could potentially compromise the system's functionality.

Metadata errors might take many different forms, but they all negatively affect resource access. The quality characteristics of metadata depend on various factors. Researchers such as Tani, Candela and Castelli (2013) and Lee (2017) concur on metadata record quality by examining various factors, namely, consistency, provenance, metadata record accuracy, and accessibility, logical coherence, timeliness, subject term specificity, and exhaustively. Lee (2017) lists additional dimensions of metadata quality as conformance, currency, intelligibility, objectiveness, presentation, and relevancy. Lee (2017) further categorises metadata quality into the following categories:

- *Intrinsic metadata quality*: Regardless of the context in which they are utilised, metadata may possess inherent validity. This means that the metadata conform to the specification of the oral history content format).
- *Contextual metadata quality*: To offer value to the intended use of the information, metadata must be timely, relevant, full, appropriate, and up to date. This implies that the metadata need to correspond with and be pertinent to the user's search query. Because of the nature of oral history material, metadata must offer choices that correspond with the user's search.
- *Representational metadata quality:* Metadata should be succinct and consistent since it evaluates how easily understood and presented the metadata is; consequently, it should be clear. This speaks to the clarity and lack of ambiguity with which metadata may be understood.

- *Representational conciseness*: The data is presented in a clear, concise, thorough, and well-formatted manner. This comprises interpretability (the technical features of the data, such as whether it is represented effectively and if the machine can handle it), versatility, and interoperability (the extent to which the format and structure of the data conform to data from other sources) (the availability of the data in different representations and in an internationalised way).
- *Accessibility metadata quality:* This refers to the metadata's accessibility and timely receipt, as well as the ease with which it can be obtained. It consists of interlinking (the degree to which entities that represent the same concept are linked to each other), availability (the degree to which data will be present, available, and ready for use), licencing (the granting of permission to reuse a data set under defined conditions), and security (the extent to which data are protected against alteration and misuse) (Sobczak 2013).

On the other hand, Király (2019) identifies the following metadata quality assessment:

(i) *Completeness:* The number of metadata elements filled out by the metadata creator in comparison to the total number of elements in the application profile.

(ii) *Accuracy*: The information in the fields relates to the resource being described in an accurate metadata record. This includes the metadata's ability to accurately represent the item being described on a semantic level.

(iii) *Consistency*: The information in the fields relates to the resource being described in an accurate metadata record. This includes the metadata's ability to accurately represent the item being described on a semantic level.

(iv) *Objectiveness*: The extent to which the metadata values give an objective description of the resource without disparaging or enhancing it.

(v) *Appropriateness*: The extent to which the values of the metadata make it easier to implement search tools on top of the repositories.

(vi) *Correctness*: How much of the language used in metadata is grammatically and syntactically correct?

The emphasis on the significance of high-quality metadata and how it might impact resource discovery, access, and utilization are the points of convergence. In the end, it is necessary to evaluate the suitability of each metadata piece in order to strike a balance between the cost of

developing the descriptions and the specificity of the knowledge that can be represented. In summary, the description and arrangement of oral history content involves creating metadata (descriptive information) about the interviews and organising them in a meaningful way that makes them accessible to users (Sobczak 2013). This process requires a metadata quality assessment to ensure metadata access for this oral history content.

2.4.5 Selection of Metadata schema

The selection of an existing metadata schema or format requires customisation to meet the needs of a specific collection and its users. Metadata schemas generally comprise semantics, content rules, and syntax (Formenton & De Souza Gracioso 2022).

Semantics refer to the metadata elements within a schema, including their names, definitions, repeatability, and whether they are mandatory, optional, or conditional. These elements describe aspects such as hardware, software, rights, preservation activities, and the lifecycle of digital objects. Content rules govern how metadata elements are selected and represented, ensuring consistency in the description of resources. Syntax provides the machine-readable structure used to encode, transmit, and exchange metadata across systems.

According to Formenton and De Souza Gracioso (2022), metadata schemas may be classified as generic or domain-specific. Generic schemas can be applied across different types of resources, whereas domain-specific schemas are designed for particular subject areas or resource types. The selection of a metadata schema should be guided by collection characteristics, resource types, and the needs of users and communities served by the collection (Gilliland 2000). Technological infrastructure and staff expertise are also important considerations for supporting access to digital collections and repositories (Simmons 2014). Furthermore, metadata interoperability remains a key requirement, although it is often challenged by locally developed metadata elements and guidelines that are not easily shared across systems (Simmons 2014; Xie & Matusiak 2016).

In essence, a metadata schema provides the structure and elements used to organise and describe metadata, whereas a metadata standard establishes the broader rules and requirements governing the creation, management, exchange, and quality of metadata. A metadata standard may incorporate a specific schema but also addresses issues such as data formats, content

requirements, and interoperability. The next section examines metadata standards used in archival institutions.

2.5 METADATA STANDARDS FOR ORAL HISTORY CONTENT

Metadata standards determine the persistence, consistency, comprehensibility, access, and representation of selected sites, collected and stored in any archives (Formenton & De Souza Gracioso 2022). Metadata standards for metadata usage and description should be adhered to, as this is necessary for the item to be found (Ramesh, Vivekavardhan & Bharathi 2015). The conceptual framework for this study intended to analyse metadata standards in such a way to determine the suitability and interoperability of metadata for oral history content. Regardless of how many versions of oral history content can be created over time, all significant properties should remain unchanged and accessible because oral history collections are available in various kinds of media and forms. To describe them, there are several diverse metadata standards. Ramesh, Vivekavardhan and Bharathi (2015) affirm that there is no one metadata standard that can be applicable to all media and document formats; a combination of them might be necessary. Therefore, it is relevant to describe the available metadata standards that can be used for oral history content. To contextualise these standards in this study, this section begins with a discussion of the principles underlying metadata standards.

2.6.1 Principles of metadata standards

Metadata creation is one of the most crucial tasks of an archivist. To achieve this, they should be guided by principles. The principles that are thought to be universal across all metadata domains and can guide the creation of any metadata schema or application (Duval, Hodgins, Sutton & Webel 2002, Greenberg 2017). Since this study aimed to propose a metadata framework for oral history content in South African archival institutions, it is appropriate to discuss the principles that guide the use of metadata standards. Metadata are a primary tool in metadata creation and access; therefore, it is valuable to understand how to integrate it into information management systems, how to create, enrich, and filter it, and, most importantly, how to manage it to allow use and access. It is significant to consider the principles that underly the use of metadata standards.

2.6.1.1 Modularity

Metadata modularity is a key organising principle for environments characterised by vastly diverse sources of content, styles of content management, and approaches to resource description (Duval et al. 2002). This is the case with oral history content that is available in a variety of formats, such as audiocassettes and other digital formats. Duval et al. (2002) emphasise that a compound schema that incorporates the capabilities of each component may be created by combining a metadata module and an instructional management metadata module. This allows modular sets to be put together to satisfy both local and domain-specific needs for a given application without unduly compromising cross-domain compatibility (Duvan et al. 2002). This suggests metadata modularity would mean dividing metadata into folders created in a way to allow grouping together related metadata, thus revealing their underlying relationships. To discover and access oral history content, metadata files should be displayed within folders corresponding to the metadata type, while showing which groups of metadata are related.

2.6.1.2 Extensibility

Metadata standards should allow extensibility to increase interoperability between metadata systems internally and externally. Metadata systems should support extensions in order to meet user search needs. While some metadata elements are likely to be unique to certain applications or domains, others should be present in the majority of metadata schemas.

2.6.1.3 Refinement

Refinement refers to the creation of metadata standards which should enable schema designers to select a level of detail suitable for a particular application (Arnold 2015). This also applies to the application domains, which vary based on the level of information required or preferred. Arnold (2015) mentions two refinements to consider, that is, the first is the insertion of qualifiers that further clarify the meaning of an element; and the second one is the definition of schemes or value sets that specify the range of values for a certain element constitutes the second type of refinements. This suggests that it may be far more helpful to know that a metadata value was created using a certain algorithm or that it was chosen from a defined vocabulary, particularly for automated processing.

Refinement is an essential principle in metadata application for interoperability purposes. Interoperability may be enhanced by the use of controlled vocabularies as it increases the precision of descriptions to increase subject access to resources (Arnold 2015). Arnold and van der Walt (2018) further indicates that the use of a particular vocabulary in each collection of metadata will allow applications to provide coherent search and browsing facilities.

2.6.1.4 Multilingualism

The translation of pertinent specifications and standard documents into many languages serves as the fundamental foundation for the advancement of a global metadata architecture (Jana & Rout 2022). Because South Africa is a multilingual nation with oral history collected in a variety of national languages, it is critical that translations be easily accessible. Adopting metadata standards that take into account language and cultural diversity is therefore essential. Both localisation and internationalisation should be taken into account. Even as internationalisation progresses, maintaining local customs might better meet the demands of any given community; character sets, common conventions, and languages are perfect for "global resource discovery" (Duvan et al. 20002; Beyene & Godwin 2018; Chelliah, 2023).

The basic starting point for promoting a global metadata architecture is to translate relevant specification and standard documents into a variety of languages (Jana & Rout 2022). South Africa is a multilingual country with oral history collected in various languages of the country; therefore, it is important that translations are made available and accessible. Therefore, it is necessary to adopt metadata standards that respect linguistic and cultural diversity. Internationalisation and localisation are equally important to consider. Supporting local customs can better serve the needs of any specific community, even while internationalisation; common rules of practice, languages, and character sets are ideally suited for "global resource discovery" (Duvan et al. 20002; Beyene & Godwin 2018; Chelliah, 2023). The current researcher noted that it might sometimes be a challenge for a metadata standard to equally support both international and local content. Thus, it is equally important to include provisions in the metadata for the description of linguistic and other cultural aspects of a resource. For example, metadata can describe the language and character set of the resource. Schellnack-Kelly (2020) agree that language could be a barrier and/or enabler in accessing archival institutions, hence oral histories. Schellnack-Kelly (2020) emphasises decolonising the archival institutions to that effect. The research concur that decolonising the archival

institutions would allow description of metadata in diverse languages. This will in turn address the language barrier in archives, since the results showed that most archival materials are in Afrikaans language and also accessible in Afrikaans.

To describe metadata, there are several diverse metadata standards. As one metadata standard cannot be applicable for all the document formats, or a combination of them may be used (Ramesh, Vivekavardhan & Bharathi 2015). In this context, this study presents an overview of the types of metadata standards and issues related to metadata creation.

2.6.2 Types of metadata standards

The objective of the study was to analyse the metadata standards used to classify oral history content. The type of metadata standard used for oral history content is an important factor in its classification. For the purpose of this study, two different types of metadata standards, which were used interdependently, are described: Standards for metadata structure provide uniform organization among entries, facilitate data searching, and allow for cross-disciplinary data sharing. Standards for hierarchical structure allow for the description of context and content (Jana & Rout 2022).

The inclusion of more specific elements and attributes in these metadata standards makes it possible to record specific details about an item that is being described. Although each of these metadata classification standards has its own unique features and benefits, they all share the goal of providing a standardised way of describing and organising archival materials. Using these standards, archivists can ensure that their descriptions are consistent, accurate, and easily searchable, making it easier for users to discover and access the archival materials they need.

Regardless of how many versions of oral history content can be created over time, all significant properties should remain unchanged and accessible. Ramesh, Vivekavardhan and Bharathi (2015) affirm that a metadata standard cannot be used to all media and document formats; therefore, a combination of them might be necessary. Thus, it is relevant to describe the classifications available that can be used for oral history. To contextualise the classification standard in this study, this section begins by unpacking what classification means in archives in general and the NFVSA in particular. Hence, the next section discusses the archival classification of oral history content.

2.6.3 Classification of oral history content

The oral history content in the NFVSA remains uncatologued and thus unclassified. The objective of this study was to analyse the metadata standards used to understand the backlog in terms of metadata classification or oral history content. Backlogs refer to materials in a repository for which the turnaround time from acquisition to access is unacceptable. Repositories have different priorities in approaching backlogs. In many repositories, technology problems keep oral histories in the back room (Mackay 2016). Da Graça Simes et al. (2016: 531) emphasise that classifications fulfil their most important goal: "to be a methodological resource for both intellectual and physical organisation and representation of information and knowledge by themes in order to promote effective and efficient retrieval." However, the classification differs depending on the organisation. For example, in archives, classifications have traditionally been related to organisation and control to obtain efficiency and effectiveness in managing records (archival classifications), which are different from other organisations such as libraries (bibliographic classifications) (Colla, Goy, Leontino, Magro & Picardi 2020).

Archival classifications are structured codes that represent organic structures, functions, and activities within records (Batley2014). They are not necessarily concepts but rather groups of records. Bibliographic classifications help organise knowledge, retrieve it, and organise records within epistemological frameworks. They promote control by institutions and allow for effective description and access. Archival and bibliographic classifications share intellectual and physical processes for organising information and knowledge, with the primary purpose being retrieval or management (Batley 2014; Tinker 2014). They were created to arrange documents for specific purposes.

Classification in archives aims to achieve recall and precision in information objects. However, access to oral history content is often limited due to lack of consensus or standardisation in classification methods, and access to oral histories remains less standardised than other materials (Wynne 2009; Raghavan & Krishnamurthy 2013). Access to oral histories remains less standardised than access to other types of materials.

Gracy and Lambert (2014) state that "No matter what metadata standard is used, the key is to pay attention to qualities such as versatility (able to ingest material from a number of different sources), extensibility (allowing for extensions to be developed), sustainability (ability to

migrate to new systems), and granularity (metadata must be sufficiently granular for all intended uses)." The nature of archival holdings fosters the assumption that it is possible to develop and apply common standards for the description of our holdings.

The Oral History Cataloguing Manual (OHCM) introduced a new approach to preserving oral histories, focusing on their context and shared provenance. Despite its significance, oral histories lack consensus and standardisation, despite their resemblance to archival materials. Despite these advancements, access to oral histories remains a challenge in archival institutions (Wynne 2009). According to Vos (2013), oral history collections are an important component of many archive collections, but there are not enough standards for description and metadata. Oral history collections are frequently left to rely on their own local standards or metadata and description guidelines outside of archives and libraries, which may not be the most appropriate for their particular requirements (Colla et al 2020).

Da Graça Simões et al. (2016) state some issues to consider when determining a local approach to classification include the physical location(s) where oral histories will be housed, the browsing behaviours of users within the finding aids, and the time available to spend choosing classification numbers for each interview. These considerations will vary from one institution to the next, depending on local preferences and the subject(s) covered by the interviews during data collection. Metadata standards are used to organise and categorise digital resources, making it easier to discover and access them. Several metadata standards have been developed for different types of resources and applications. The focus of this study was on metadata standards developed for use in archival institutions. Therefore, it was relevant to describe metadata standards in relation to oral history content.

2.6.4 Metadata standards and oral history content

Metadata standards provide consistent and rich metadata for long-term digital resource discovery, use, and integrity. They facilitate effective searching, improved digital curation, and formalize element structure, despite the lack of a definitive, unique metadata standard for diverse oral history content formats and institutions (Bountouri 2017). There are several metadata standards, but this study selected a few based on their relevance and use in the public archives of South Africa, where oral histories are also collected and organised. Only the metadata standards underwritten by ICA are explained.

Commonly used archival standards include the ISAD(G), which specifies criteria for archival description, while EAD normalises the electronic exchange of data between archives. Other standards include the International Standard for Describing Institutions with Archival Holdings (ISDIAH), which enumerates elements of description and coding information about archives; International Standard Archival Authority Record (ISAAR), which performs similar functions but about archival records; and Preservation Metadata: Implementation Strategies (PREMIS), which allows metadata coding for digitised objects (Sobczak 2013).

(a) International archival descriptive standards (general)

The ISAD(G) and ISAAR (CPF) were published by the International Council on Archives (ICA) (1999; 2004). The ISAD(G) is a set of guidelines (developed by the International Council on Archives (ICA) to standardise the description of archival materials. Provides a framework for organising and describing archival records, making it easier for archivists and researchers to find and access them. The ISAD(G) consists of seven main sections (description includes 26 elements in seven blocks), each providing guidance on different aspects of archival description, namely, identity of the archival material, context of the archival material, content and structure of the archival material, conditions of access and use of the archival material, sources of information about the archival material, rules or conventions, and notes.

The ISAD(G) specifies a multilevel structure for archival record descriptions, including key details like reference code, title, creator, data, description level, size, and carrier, for efficient management and sharing. This information allows for managing archival records and sharing them with users. Overall, the ISAD(G) is a comprehensive and flexible framework that can be adapted to different types of archival material and different local contexts. By following the guidelines outlined in ISAD(G), archivists can ensure that their descriptions are consistent, clear, and informative, making it easier for researchers to discover and access the records they need.

(b) International Standard Archival Authority Record

Similar to the ISAD(G), International Standard Archival Authority Record (ISAAR) contains a list of fields describing the items, and has 32 elements in total, split into six parts.

(c) Encoded Archival Description

Encoded Archival Description (EAD) is a metadata standard that allows for the encoding of finding aids, which are inventories of archival collections. EAD includes elements such as title, creator, date, scope, and content. EAD was adopted by the SAA in 1999 and allows archival metadata to be distributed electronically (Nimer 2010). It is commonly asserted that because EAD presents data to users in a more consistent and easily shareable manner, it can enable more productive reference interactions. The EAD technical structure standard provides serialization capabilities for data exchange and persistent storage, and it is entirely compatible with the ISAD(G) descriptive standard.

(d) Preservation Metadata: Implementation Strategies

Preservation Metadata: Implementation Strategies is the international standard for metadata that supports digital item preservation and guarantees long-term usability. PreMIS support is included into a number of open-source and commercial digital preservation tools and systems, and it is used in digital preservation projects all around the world. Everyone uses oral histories as valuable primary sources, including researchers, students, family members, and genealogists (Ritchie 2011). Some oral histories are now digital, but many still exist in previous formats such as minidiscs, mini-DV, and audiocassette tapes. Many times, oral histories contain multiple components, including audio, indexes, transcripts, and sometimes additional materials, such as photographs or handouts. It is therefore necessary for information providers to, as Boyd (2011) argues that one should ensure that each of the component of a collections' parts functions as a whole in a user-friendly, effective, practical, and intuitive way; that they can be more readily interacted with and discovered; that they can be distributed more broadly and effectively; and that they are conserved more responsibly. However, there have been major advances in standardising methods for oral history content. Metadata standards ensure better access to oral history. Therefore, the next section focuses on access to oral history.

2.7 ACCESS TO ORAL HISTORY CONTENT

Oral history helps identify, appreciate, and respect various aspects of heritage (SAHO 2019). The objective of this study was to determine access to oral history content in the NFVSA. Providing adequate access to oral history content is an essential function of any archival institution. However, oral histories deserve attention because they are unique to a particular institution and often cover topics or viewpoints that may be under-represented in published works (Wynne 2009). Providing access is essentially about enabling interested parties to access

the information they seek. The two most important factors in achieving this are providing access to context information about archives, that is, archival description, and access to the archives themselves (Edmondson 2016). Edmondson (2016) further states that planning access to records involves considering material type, rights, licenses, user access, technological advancements, user attitudes, services, infrastructure costs, reuse, sharing, and Web 2.0 solutions. This suggests that one of the factors to access oral history content is the formats in which they are. The next section explains various formats of oral history content.

2.7.1 Formats of oral history content

The nature of archives is 'mixed or hybrid' as they include records that are digitised and records that are born digital, (Sobczak 2013; Jaillant & Caputo 2022). Some oral histories are now born digital, but many still exist in previous formats such as audiocassette tapes, minidiscs, and mini-DV (Bilushchak & Fedushko 2016). For this reason, there is an increasing need for digitisation in archival institutions. One approach to providing access to oral history collections is to make them available through online platforms or digital archives. For example, oral history collections should be accessible through their online metadata and search functions. Similarly, the NFVSA provides access to its oral history collections through a searchable database of oral history recordings and transcripts.

Boyd (2013) states that archivists tend to treat the different components of the oral history information package (audio, video, transcript, index, and metadata) as separate entities. Users of typical archival platforms can access an audio record, access a transcript, and even search a text while listening to or watching the interview. However, the different components are rarely joined into one integrated experience that can help access and use.

In general, providing access to oral history collections requires careful consideration of a variety of factors, including the ethical and legal implications of sharing these materials. The next section focuses on the components in access to oral history content.

2.7.2 Components on access to oral history content

In 2020, the Oral History Association Metadata Task Force (OHAMTF investigated how oral history practitioners describe their collections to allow discovery and accessibility for the public. These categories are directly linked to steps in a metadata workflow, including file

naming, capturing, processing, quality control, production tracking, searching and retrieval, designing, and storage (Cherry & Mukunda 2015).

Access to oral history includes the following key components to be considered:

(a) Permissions and ethical use

Bradley and Puri (2016) indicate that although searchable online systems make it possible to access primary sources, including content from oral histories, they further complicate already complex ethical dilemmas. The challenges to accessing oral history relate to the intellectual property and ethical issues of individuals. Oral history collections provide unique and valuable information on the lives, experiences, and perspectives of individuals and communities. However, access to these collections can be challenging due to issues such as confidentiality, copyright, and ethical considerations. For example, confidentiality and consent should be considered to ensure that there are no risks that individuals participating in oral history interviews, are affected. Furthermore, issues of copyright and ownership can be complex, particularly if multiple individuals or organisations have a stake in the oral history materials.

Wood,Carbone, Cifor, Gilliland & Punzalan (2014) indicate that the risk associated with metadata is that it may reveal susceptible persons and populations through linked descriptions, excessively thorough authority files, and subject descriptors. For instance, the use of connected data to specific historical materials, such letters and photos, enables complex data visualisation and links within and between repositories (Gilliland & Willer 2014). Hence, there is still no clear response to the question "the role and scope of, as well as who has jurisdiction over such re-description or alternative descriptions of oral history resources", (Dougherty & Simpson 2012).

It is thus important for archivists to consider ethical obligations and verify if files can be shared; whether there are ethical or legal restrictions that limit curation, sharing, and preservation; and whether there is adequate documentation to that effect. This also has legal implications. The study conducted by Schellnack-Kelly (2017) found that the obligations and rights that historians and archivists have with regard to following moral principles when collecting oral histories and sharing information about them.

In general, archives are not free to publicly exhibit, perform, or exploit material in their collections without approval of the relevant copyright holders. For example, MacKay (2016)

adds that archival institutions should respect copyright rights and maintain the rule of law. However, as rights are sold, resold, and creators die, it becomes increasingly difficult to determine ownership, leading to an increasing number of 'orphan' works (Edmondson 2016; MacKay 2016).

Archival access operates within a framework of contractual and copyright law. Access and, to some extent preservation, activities are governed and constrained by the legal entitlements of copyright holders. Edmondson (2016:81) summarises the legal implications relating to access to oral history content as follows: “It becomes the obligation of archival institutions to guard the interest of their stakeholders by ensuring respect for intellectual property while also their obligation to serve the public interest by providing access to their collections.”

(b) Metadata

The archivist should determine what metadata schema is used and ensure that all files and documentation are at least using a constant standard that will aid discovery and indicate who can access and reuse files. With respect to documentation, they should determine whether the documentation of the data is sufficient for a user with similar qualifications as the author to understand and reuse the data. If not, additional documentation should be recommended or created. This is important to ensure that high-quality metadata are created.

(c) Technical evaluation

An inventory of the recordings, transcriptions, annotations, and administrative data that archivists want to curate should be compiled. In order to ascertain whether the papers are in open format or proprietary, they will also look for quality assurance and usability problems, such as missing transcriptions, administrative documents, timestamps, and translations. This will help identify the software needed for accessing and interpreting oral history files. As suitably indicated by the foregoing stages, inventory and assessment files are completed for preservation purposes, such as y-checks, and are ready for long-term storage and access.

(d) Transcription

By enabling viewers to conduct a comprehensive text search on an interview, transcripts and video captions contribute to the discoverability and accessibility of interviews.

(e) Public access and discoverability

Access to oral histories and metadata should be granted in accordance with the collecting institution's guidelines and/or in accordance with the abilities of the participants (Kata, Milbrodt, Sielaff & Vos 2020).

OHASA (2019) specifies some of the issues relating to access to oral history content. The current study recognises the distinct nature of the metadata descriptions that are done in archival institutions, unlike other information organisations (Hider 2015). Thus, it is relevant for this study to propose a metadata framework as a guide for South African archives. The proposed framework is discussed in chapter six. The challenge in this study is to identify specific metadata for oral history content, as a multiplicity of metadata standards for describing content exists. Embedded metadata allows data to travel with the file and is very useful in cases in which records change location or are reused and searchable in an immediate way. Hence, the next section focuses on technologies used for oral history content.

2.8 TECHNOLOGIES USED FOR METADATA OF ORAL HISTORY CONTENT

The objective of this study was to determine technologies used for oral history content in the NFVSA. Technologies play a key role in how different organisations provide their services. Archival institutions are affected by their access to archival materials. The conceptual framework of this study formed the basis for determining the technologies that are applicable to metadata and access to oral history content. Thus, in this study, it is relevant to understand the technologies used for metadata around providing access to oral history content. The most obvious technology with which technology archivists working with oral history content interact is the information management system.

2.8.1 Information management systems in archival institutions

The information management system plays a crucial function in any organisation, including archival institutions. An information management system refers to any framework of software that facilitates the collection, storage, organisation, and distribution of information (Ladley 2012). Archival institutions require an information management system to create, store, and make available oral history content to the South African public. The metadata system is crucial for metadata creation and access.

2.8.1.1 Metadata system

The metadata system has transformed many organisations, including archival institutions and their oral history metadata access. Therefore, it is important to explain the metadata system that would be relevant and assist in metadata creation for oral history content.

Poynton (2012) suggests the following guidelines for successful implementation of a metadata system.

- Clear guidelines for what metadata should be encoded and how to interpret metadata during decoding should be incorporated into the metadata system's thorough development. This means that the system should be as simple as possible.
- The metadata system's functionality should never be compromised by the addition of metadata thanks to careful design of the metadata system.For example, the system should not duplicate existing available data on the system. This should be done to minimise the possibility of conflicting metadata records.
- The metadata system should allow for addition as an augmentation. The principle of enrichment proposes that the system should allow the addition of keyword search to ensure discoverability of, and access to, information (Alemu 2015).

The metadata system is considered to be the best way to minimise the risk that digital objects becoming not accessible (Sambo, Urhefe & Ejitagha 2017). A good metadata system should allow users to quickly and accurately discover and access information and/or oral history content so that they can make informed decisions. This suggests that metadata should also be managed accurately and timeously. Hence, metadata management is important in information management of any organisation, including archival institutions.

2.8.1.2 Metadata management

The usefulness of the metadata depends on the management of the metadata. Metadata management processes are used to determine and evaluate how each data connection can be drawn from an authoritative source. Loshin (2013) believes that in metadata management, authoritative sources include features that aim to maintain ‘maximum productivity and efficiency.’ Allen and Cervo (2015) add that these features include data security, data governance, and data integration.

First, data security involves strategies to protect theft and/or unauthorised access, use, or destruction of sensitive data (Allen & Cervo 2015). Given the sensitive nature of some of the oral history content, security is crucial to managing metadata and access. This is in line with the best practice of OHA (2019), in which permissions and ethics are the core components of access to oral history content. Metadata management should include security controls and risk management to protect sensitive data (Allen & Cervo 2015). These involve a range of policies; hence, data governance remains an important issue in metadata management.

Data governance involves tracking and managing data to ensure that it is accurate and up to date (Borek, Parlikad, Webb & Woodwall 2014). This would suggest that metadata are monitored and evaluated to ensure that they meet the requirements for successful discovery and access to oral history content. Additionally, this would also help keep track of who accesses the oral history content and manage what users could do with it. This supports access to oral history content and the benefits associated with access. As already stated, the complex nature of oral history sometimes makes it difficult for metadata creation, thus hindering access. It would require a system to assist in metadata creation and access. As such, data integration plays an important role in a metadata management system.

Lastly, metadata management involves data integration, which creates a common repository for data that many applications may access and share. (Loshin 2013; Allen & Cervo 2015). To make sure that the data is compatible with the programmes that will use it, the system can include tools for "data cleansing and transformation," (Allen & Cervo 2015). For oral history content in different formats, data integration would assist in linking and connecting all the formats that belong together, but also with other related data. Related to metadata management, the intent is to capture oral history content in the metadata repository, as well as the metadata associated with it. Thus, for the purpose of this study, it is necessary to discuss metadata repository as part of the information management system.

2.8.1.3 Metadata repositories

The objective of this study was to analyse technologies used for oral history content; therefore, it is relevant to discuss the metadata repository in relation to access to oral history. The metadata repository stores metadata. According to Reeves (2013), the metadata repository is a tool for modelling, profiling, transforming, and managing the relevant metadata for the development and operation of data integration solutions (Plotkin 2021). Archival institutions

benefit from metadata repositories, as they link the metadata from the various tools to each, thus offering consolidated metadata, and between various formats as with oral history content. Plotkin (2021) further emphasises that a metadata repository should have a description, which includes a schema, view, dimensions, hierarchies, as well as data locations and contents. This suggests that a repository includes databases and their contents that are put together for authorisation and control to access of oral history content. If the metadata repository is well organised, it will improve the discoverability and accessibility of oral history content in archival institutions.

Metadata repositories typically hold reformatted digital content (Lambert & Frisch 2013), for instance, Microsoft Word files that have been converted to PDF format for online display. The archival value is typically found in the original digital object, which should be kept as the master in order to track all modifications, interactions, and the preservation of the linkages between the many documents (Matusiak, Tyler, Newton & Polepeddi 2017). Archival document backup copies can be kept on reliable external hard drives or on DVDs that are kept in acid-free compartments and kept at a controlled temperature. But the technology and services available to support collections will determine how long this storageware lasts (Sandaver 2020). On the other hand, a metadata repository remains significant, and it would thus require a database and finding aids containing metadata to support the discoverability and access to oral history content.

2.8.2 Databases and finding aids in archival institutions

The main aim of any information system is to transform data into information that can be accessed and used. A database is such a tool that helps to achieve this goal. A database is an organised collection because in a database, all data are described and associated with other data (Bourgeois, Smith, Wang, & Mortati 2019). Archivists are increasingly turning to databases as a way to manage their collections and make them accessible. Databases and metadata that provide users with familiar forms and methods for locating content are necessary in relation to the accessibility of oral history content. In addition, in the NFVSA, a database is used to access archival materials, including oral history content. It should be possible to search databases by metadata in various categories to access archival materials. Therefore, when creating a database, one should not simply apply an overall approach to finding aids.

Finding aids are tools for retrieving and accessing collections in an archive. For this reason, when creating finding aids, archivists should focus attention on own needs and on reproducing the arrangement of the record instead of the needs and interests of users (Sobczak 2013). Sobcaak (2013) adds that the finding aids should allow users to access oral history content. At the NFVSA, the NAAIRS is used for this purpose (NARSA 20216). The NAAIRS is a finding aid for the NFVSA and it provides an important reference for the retrieval of the actual archival material in a repository.

This study focused on how technologies can be used to access oral history content and the metadata creation process within these technologies. The next section focuses on software applications in archival institutions.

2.8.3 Software technologies in archival institutions

Software used in archival institutions to describe records allows virtual arrangement and archival description (Király2019). The software's electronic format gives it advantages over other locating aids, such as almost limitless data collecting and the ability to be updated and augmented with new requirements as they arise, add Király (2019). The conceptual framework of this study focused on technologies intended to enhance metadata access and to supplement ways of accessing oral history content with the enriched metadata. It is thus important that the software selection should be made carefully, taking its purpose into consideration (Feliciati & Alfier 2014). When selecting software, one should consider a number of requirements to ensure its efficiency and functionality (Borek et al. 2014). The fundamental requirements that must be taken into account during the design of a software, according to Király (2019), are as follows: the abstraction of the metadata schema, how to address distinct parts within metadata records, the measurement workflow, a common and powerful interface for the individual metrics, and interoperability. Király(2019) further indicates that software for the processing of records should involve managing archives, creating fonds structures, describing fonds, classifying, assigning reference codes, selecting record description forms, building series, processing levels, indicating date ranges, specifying storage periods, and sorting records.

This is true with the NFVSA because, as a national archival institution, it should manage metadata from the provincial archives in the country. Thus, it will need software that describes and generates metadata to maximise access to oral history content and interoperability. The software required at the NFVSA should be scalable to handle a large number of metadata

records in a fair amount of time (Király 2019:5). This led to organisations accessing their collections through web-based platforms. Archival institutions should also consider the web-based system for metadata access and oral history content.

2.8.3.1 Web-based system for oral history content

In the digital age, users are moving to the use and exploration of web-based platforms as the new normal practice for accessing archives and interacting remotely with special collections (Mosako & Ngoepe 2021). The presence on the web can be an effective way to provide access to oral histories (Bilushchak & Fedushko, 2016). Bilushchak and Fedushko (2016) add that the benefits of the web for oral histories include global reach, hyperlinks to related materials, tied retrieval capabilities and access privileges. It is very important to provide a web-based system to enable users to access the oral history content they are searching for.

The importance to search using a web-based system of an online archive (by simply typing in the name of an organisation that owns a website), perform an advanced search using metadata (by title, person, or subject area, for example), or browse them (by record type), cannot be overstated (Boyd 2013). The OHMS is one of many open source web-based tools available, but it is cost-effective, user-friendly, and can be used for a wide variety of purposes (Royles 2016). However, regardless of the repository being used, the OHMS has been built to allow the user to personalize the system and enhance the experience. It was intended for the OHMS to use transcribed oral histories (Boyd 2013). This will require a search system that will allow access to oral history content.

2.8.3.2 Search system and archiving

Archival institutions rely on the search system for access to their collections and archiving. Thus, it is important to pay careful consideration to the search system to be adopted and implemented. Fraser and Halvey (2017) discuss some recommendations for designing an oral history search system:

- Implement a number of engagement platforms and make sure that the search system or archive can be accessed by a variety of user groups, including the public, researchers, younger audiences, and people with special needs.
- To encourage more participation, make the technology and collections available through workshops, online manuals, tutorials, institutional partnerships, and outreach

initiatives. This could take the kind of community outreach initiatives, university seminars, or school visits.

- To protect interviewees and organizations, clearly define your ethical and consent rules. In order to monitor and share materials, organizations should also work to establish a central repository of materials that can be efficiently managed and monitored. This might entail, among other things, offering restricted access to specific areas, options to restrict access for predetermined times of day, and worldwide access to information.
- Create podcasts, video segments, and mobile applications for the search system to broaden its audience and boost user engagement.
- Prioritise keyword searches, tags, and metadata to make sure users from diverse backgrounds can locate the exact phrases or information they need in collections.

Fraser and Halvey (2017) further add that transcripts that are helpful for research and during the early phases of the search process, along with keywords and metadata with special attention to tags, filters, metadata, and multilingual search terms, are the most crucial components of any oral history preservation and search system. In addition to addressing accessibility challenges, legal and ethical issues, and user attractiveness, these additions would boost engagement. Mobile applications, video segments, online tutorials, and consent policies are crucial for metadata access for oral history content. This would mean that the search system should transform archival institutions and access to oral history metadata. In the labor-intensive process of developing metadata for oral history content, people generated the majority of the resource finding information prior to creation. This is because metadata creation for oral history content is also labour intensive when performed by human elements. Hence, artificial intelligence (AI) may be recommended for metadata creation as a way to cap the hard labour relating to that. Therefore, it is important to discuss artificial intelligence and its relevance in metadata creation.

2.8.4 Artificial intelligence and metadata creation

Artificial Intelligence (AI) has changed the archival environment. Corrado (2021) provides a basic overview of how AI can be used to extract metadata from various types of files. This would be the case with the oral history content. For example, language processing can extract meaning from speech in audio files by combining various techniques and extra tags from video files. This is done to reduce the labour-intensive nature of metadata creation for oral history

content. This suggests that the creation of metadata can be automated to some extent for easier and less effort to provide oral history content.

By combining description, classification, and organisation to assist create more effective information repositories, metadata are an essential component of the information architecture needed to help bring order to the chaos of the web (Alma'aitah, Talib & Osman 2020).These authors emphasise the importance of enhancing metadata access for heritage and cultural collections in a digital age. Like the resources themselves, the sources of metadata will vary in quality and be arranged for distinct purposes to represent the various goals and economic models of information providers. There is no question that unforeseen possibilities and niches will be created by the societal policies, organisational priorities, and commercial pressures that shape the online information spaces (Feliciati & Alfier 2017). There will need to be some convergence of encoding formats and widely accepted semantics in order for these prospects to be realised.

However, archival institutions have turned finding aids into the partially integrated electronic record management systems. Accessibility, electronic records management, and scanning systems are growing in popularity, as such AI remains a necessity for the current and future metadata practices (Giovanni, Blanke, Jeurgens & Noordegraaf 2021).

There are also technologies that help enhance queries and are discussed in the literature. Only mentioning these is appropriate here primarily because they make use of classificatory tools. Some of the widely employed technologies are the following (Raghavan & Krishnamurthy 2013: 312):

- Suggesting related queries (by search engines)
- Relevance feedback: marking the relevant documents in the initial set and using this to reformulate the query
- Using lexical tools for query expansion/reformulation (for example, Thesaurus or WordNet)

However, using technologies in archival institutions is not without challenges. The risk of prolonged accessibility and long-term preservation of digital assets increases with technological obsolescence (Zierau & Peyrard 2016; Masenya 2018). It has been assumed for far too long that corporate and policy makers are unaware of the historical relevance and worth

of digital heritage (Corrado & Moulaison 2014). Most digital objects are poorly organised or lack adequate descriptions, which makes it more difficult to provide long-term access to digital information (Ross & Hedstrom, 2005). According to Styblinska (2006), maintaining continued access necessitates keeping up with technical advancements and transferring digital artifacts from antiquated to modern operating systems, file formats, and storage media.

Metadata are the best course of action for handling software obsolescence, hardware malfunctions, and backup concerns (Sambo, Urhefe & Ejitagha 201 7). This is because well-maintained metadata identifies data was created and when it has to be disposed, thus obsolescence could be determined and addressed. According to Bountouri (2017), metadata is information about items contained in collections and their unique uses, such as object identification, search, retrieval, and collection management. Trust is an issue in today's world of ever evolving technologies, and audit trails, logs, and metadata are more direct means of establishing trust (Duranti, Rogers, McLeod & Shepherd 2019). According to ICA (2016), digital records must be dependable, legitimate, and available for as long as needed. The integrity and reliability of digital records are primarily determined by metadata systems.

The intersection of special collections and information technologies is how metadata is created. The metadata specialist must be completely aware of the object they are describing in order to accurately describe it (Sandaver 2020:38). To assess and confirm the reliability and accessibility of oral history content, an appropriate technological infrastructure that complies with metadata and other international standards is also necessary. Due to improper organisation or lack of suitable description, the majority of digital objects make it more difficult to provide long-term access to digital information (Ross & Hedstrom 2005; Masenya & Ngulube 2021).

Metadata remain the most successful and standardised way for discovery, but they are costly and impractical for many materials available online. Although metadata continues to be the most effective standard for finding, numerous online materials make them expensive and unfeasible. Metadata also offers a surprisingly affordable resource finding option for a wide range of information demands. As a result, AI would enable web search engines harvest and index a large portion of the internet and provide low-cost index access to it (Jaillant & Caputo 2022).

Oral history content is being preserved, accessed, and used more effectively through modern technologies. On the other hand, technology utilisation can provide certain difficulties for oral

history curators. It necessitates having access to a wide range of standards and best practices for the creation, gathering, and use of metadata in addition to technical know-how, analytical tools, and procedures from other disciplines (Fraser & Halvey 2017). Furthermore, it is critical that descriptions take into account the systems perspective, which encompasses all facets of locating, contextualising, distributing, retrieving, utilising, and recycling these assets (Boyd & Larson, 2014). However, new technical advancements have put a heavy burden on resources and present a number of threats to the narrator and archive (Boyd & Larson, 2014; Larson & Janakiraman 2019).

The research thus assumes that everyone uses oral history content as valuable primary sources, including researchers, students, and family members. One may also conclude that multiple elements are frequently included in oral history content, such as audio, transcripts, indexes, and occasionally extra resources like handouts or photos. Therefore, as argued by Boyd (2011), information providers must ensure that the collections' constituent parts function together in a way that is both user-friendly and intuitive, as well as that they are more broadly and effectively distributed and responsibly conserved.

A hybrid approach involves local and central metadata storage, with local formats being richer and central formats being simple for interoperability. Interest in this approach is centred on the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), which allows distributed metadata to be centralized. Modern technologies have transformed archives into virtual organizations, managing digital native and digitized content, and enabling national access to all archival holdings (Sobczak 2013).

According to Sobczak (2013), archival information without its origin context is valuable for archivists and users. Acquiring data from archival materials is easier for documents with identification elements. For historical objects without clues, researchers formulate hypotheses based on speculation. Using specialised tools and early data management can improve preservation and create a more precise information context (Ritchie 2011).

One may understand that the previous systems for collecting information on archival holdings involved filtering out specific pieces of information from a large amount of material, which was time-consuming and failed due to the large number of records and the lack of prediction of user queries and the massive amounts of data needed for future generations. This is because the method used for describing units involved punched cards and finding aids, allowing

correlation-based information search (Miller 2022). However, the idea of creating records allowed for standard description and limited vocabulary, aiding in searching for records.

Metadata schema vary according to the type and quantity of data items, the identification of required fields, the encoding specifications, and the application of data content and value standards vary throughout metadata schemas (Miller 2022). As such, the choice of schema has an impact on the depth of the description. When creating a customised metadata profile or adopting a schema, there are a number of things to take into account (Turner 2017; Miller 2022).

Kennedy (2008) provides a guide for professionals to choose a metadata schema, considering factors like users' needs, metadata staff expertise, resources, access to digital collections, relationships, collection scope, metadata harvesting, interoperability, and quality control. The guide emphasises the importance of adopting a schema that enhances access and retrievability of oral history content, requiring a practical approach (Formenton and De Souza Gracioso 2022) Without an effective metadata system, oral history content may neither be in context nor be described in a standardised way. Therefore, the OHA (2019) list gives the principles and best practices for oral history:

- Interviewers, organisations that sponsor interviews, and organizations in charge of maintaining oral history interviews should be aware that original recordings need to be properly cared for and stored from the moment they are made.
- Interviewers should take care of the metadata by recording their procedures and preparation, including the specifics of the interviews, and sending that data to the repository that will be handling the interview's preservation and accessibility.
- In order to establish the status of an interview, collection of interviews, or interview project as a "trusted body" of knowledge, it is necessary to capture the context of the interview. This is where curation and care intersect.
- It entails taking care of oral history interviews in all of their manifestations, both original and derived, from the time of their inception (or even earlier) into the infinite future.

- Interviewers, organizations that sponsor interviews, and organizations in charge of maintaining oral history interviews should be aware that original recordings need to be properly cared for and stored from the moment they are made.
- Interviewers should take care of the meta-data by recording their procedures and preparation, including the specifics of the interviews, and sending that data to the repository that will be handling the interview's preservation and accessibility.
- In order to establish the status of an interview, collection of interviews, or interview project as a "trusted body" of knowledge, it is necessary to capture the context of the interview. This is where curation and care intersect.
- It entails taking care of oral history interviews in all of their manifestations, both original and derived, from the time of their inception (or even earlier) into the infinite future.

Broadly understood, metadata make it possible to discover themes and meaningful relationships within oral history content and other digitally represented resources. Therefore, the current researcher developed a metadata framework in accordance with the OHA principles and best practices.

2.8.5 Digitisation of oral history content

Digitisation refers to creating a digital representation of physical objects or attributes such as scanning a paper document and saving it as a digital document (Note 2020). Simply put, digitisation refers to converting items that are non-digital into a digital representation thereof. Digitisation of oral history collections can provide many benefits, such as increased accessibility, improved preservation, and easier sharing of these important historical resources. However, it is important to carefully plan and manage the digitisation process to ensure that the digital files are of good quality and can be preserved for future generations (Dobreski, Park, Leathers & Qin 2020). The digitisation process involves several steps, including selecting the recordings for digitisation, preparing the recordings for digitisation, digitising the recordings, and creating metadata to describe the digital files.

The following are some important considerations when digitising oral history collections (Eíto-Brun 2015; Sigauke 2017; Note 2020):

(a) *Selecting the recordings*: When selecting recordings for digitisation, it is important to consider factors such as the quality of the recordings, their historical significance, and their potential value to researchers and other users.
(b) *Preparing the recordings*: Before digitising the recordings, they should be inspected to ensure that they are in good condition and that they can be played back properly. Any necessary repairs or cleaning should be done before digitisation.
(c) *Digitising the recordings*: The digitisation process involves transferring the analogue recordings to digital formats. This can be done using specialised equipment such as analogue-to-digital converters, which convert the analogue signal into a digital signal that can be stored on a computer or other digital storage device.
(d) *Creating metadata*: Metadata are information that describes the digital files, such as the date of the recording, the names of the interviewee and interviewer, and keywords or subject terms. Creating metadata is important for organising and describing digital files and can make it easier for researchers and other users to discover and access oral history materials.
(e) *Preserving the digital files*: Once the oral history recordings have been digitised, it is important to ensure that the digital files are properly stored and preserved. This includes backing up files and storing them in secure, climate-controlled environments to prevent damage or loss.

Although digitising this content helps with the issue of physical access to archives, users still need to be able to locate the information they need. Analog content requires specialised software to process, while digital content is processed nearly entirely automatically due to electronic records management systems that automatically generate the required metadata (Zierau & Peyrard 2016). Digitisation would sometimes involve the migration of metadata and materials. The next section focuses on migration of oral history content.

2.8.6 Migration of metadata for oral history

Technologies are not only popular but can also provide effective access to oral histories (Wynne 2009). Metadata should be managed in a system that facilitates access and is interoperable with other systems when future migration of metadata is necessary. Due to incompatible system used for metadata, sometimes the entries should be migrated to the newly

adopted system. Migration should be done to ensure users with search capability to oral history both in analogue and digital formats (Boyd 2013:96).

There are three methods for moving data from an analogue to an electronic environment. If previously used finding aids are not created in compliance with current information quality standards, or are unreadable, missing, damaged, or otherwise worthless, the only viable approach is to describe the fonds from scratch. The second method entails retyping current finding aids by hand, while the third and most widely used is solution uses optical character recognition (OCR) software (Edmondson 2016). The NFVSA requires more of a format migration. Format migration is the process of converting/transforming a file from one format to another. In this way these transformed formats should be made available and shared through various spaces. It is also evident that archival materials, including oral history content, should be made available through shared spaces that are not limited to a single institution but to a global audience through services (Mosako & Ngoepe, 2021). Such a shared space would greatly simplify the process of archiving and offer higher quality electronic materials.

Migration of metadata for oral history content may pose some technical and metadata-related challenges (Gill, Vos,Villa-Torres & Ramirez 2019). For example, the technical issue may be that there was not enough computer RAM to multitask during the entire migration process (Duvan et al. 2012; Edmondson 2016). This may result in some individual record files not being uploaded, as such the entire record needing to be restarted upon occurrence of the error. For this reason, metadata and file uploads could not occur at the same time.

Secondly, the challenge could arise from not being able to work on several records at once because the computer might move slowly between jobs in several programmes. The study conducted by Edmondson (2016) found that migration of oral history content metadata computers was very slow, as oral histories, whether converted from analogue or born digital, need audio processing, which required computers with sufficient memory. The mapping of records from one schema also posed a challenge. Edmondson (2016) recommended that archival institutions should consider including separate dedicated workstations in their budget.

Digital information objects that are currently being created must have metadata that allows them to exist independently of the system that is currently being used to store and retrieve them if they are to have any chance of surviving migrations through successive generations of

computer hardware and software or moving to entirely new delivery systems (Groenewald & Breytenbach 2011; Gilliland-Swetland 2000).

This study was aimed at proposing a framework as a solution to the challenges associated with access to metadata for oral history content in South Africa.

2.9 FRAMEWORK FOR METADATA FOR ORAL HISTORY CONTENT

Metadata schemas differ in the number of data elements, mandatory fields, the encoding requirements, and the use of data content and value standards (Xie & Matusiak, 2016). The challenge of creating and/or choosing a metadata schema that is best suited for metadata creation has been around for a while (Haynes 2018). Haynes further indicates that the decision on the choice of metadata schema has an impact on the extent and type of description. The researcher concurs that significant thought must go into choosing the optimal metadata format for oral history material in order to produce high-quality metadata that promotes discoverability and accessibility. As a result, several factors will affect the adoption and/or development of a customised schema.

When choosing a metadata schema, the focus should be on potential users and their needs, collection scope, expertise of metadata staff, time and financial resources, type of access to a digital collection, relationships to other collections, metadata harvesting, interoperability, and level of maintenance and quality control (Kennedy 2008; Miller 2022). Miller (2022) adds that the metadata creation process begins with an assessment of the nature of the collection, as well as cooperative and institutional information systems. This is crucial to selecting the metadata schema that will fit the unique nature of oral history content, which requires unique metadata creation processes. It is thus critical, according to Miller (2022), to identify a single unique metadata schema and standard to provide an information infrastructure with coherent methods of description and access that transcend historical boundaries of countries, cultures and languages (Ramesh, Vivekavardhan & Bharathi 2015).

Ngoepe (2020:310) states that “oral history constitutes an important aspect of any archival institution and forms a major component of society’s culture and history.” As such, oral history content requires a description beyond the boundaries of language and culture. Language plays a major role in local standards since descriptions are frequently adapted to users' expectations, and language is then adapted to those users' expectations (Knight 2022). According to the

researcher for this study, one should start with the language issue in order to create a metadata schema that takes into account the diversity of different languages in the archives. Farnel, Shiri, Campbell, Cockney, Rathi and Stobbs (2017) emphasise that it is crucial to adopt metadata schema that respect cultural and linguistic. Consequently, for effective cross-language retrieval, multilingual metadata descriptions must be present (Király 2019). As such, archival institutions should adopt a metadata schema that embraces the multilingual nature of oral history content. However, without an effective metadata system, oral history content may neither be in context nor be described in a standardised way to allow for the multiplicity of these languages.

Archives need to use technologies in their metadata processes to allow optimal access to, and use of, oral history content. For example, the study by Bilushchak and Fedushko (2016) investigated the use of technologies in museum and archival activities in Ukraine. The study concluded that the use of ICTs in the creation of electronic, digital, and multimedia archives solves several issues, such as that access is given to information, regardless of time, place, space, or social origin. Therefore, the metadata schema requires a practical approach to the description process to provide consistent access points to the oral history content, increasing access and retrievability (OHA 2019). To this effect, OHA (2019) proposes a metadata framework that consists of five major components namely:

- A schema addresses the components and types of data that should be selected for recording.
- a list of terms selected specifically for the designated categories along with their respective definitions.
- conceptual model framework, which explains the relationships between various information concepts.
- Metadata schema categories are a standard and practical approach to capture specialized information (vocabularies, for example) (for example, cataloguing cultural objects).
- Encoding pertains to the presentation of the metadata (for example, XML).

The researcher understood that, in line with the OHA principles (2019), it is important to adopt a metadata schema that would increase access to oral history content in whatever format is available in archival institutions and beyond. For this reason, the researcher, in recognition of

these principles, proposes a framework that can be used in South Africa to meet the unique nature of oral history content and the challenges related to its metadata creation. This is done with knowledge of the various factors to be considered in selecting a metadata schema as well as the digital environment in which archival institutions operate. Additionally, the framework for the process of describing oral history can be labour-intensive, especially when it involves accurately documenting the fonds or materials of the objects such as recording object numbers as well as the content type, size, materials, and extent of each item (Király 2019). To gain a better understanding of proposing a metadata framework that could guide metadata schema and metadata creation for oral history content, it is relevant to focus the next section on the related studies that have been conducted on the topic.

2.10 SUMMARY

This chapter reviewed the literature on the metadata schema of oral history content. The conceptual framework developed for this study provided guidance and was linked to the review of the literature. The chapter began by describing the metadata schema used in archival institutions. These descriptions covered the types of metadata and valuable metadata elements for the description and access of oral history content. The chapter ended with an explanation of metadata quality.

Following this, the chapter delved into metadata standards that provide guidelines for creating metadata for archival materials, including oral history content. It explained the metadata principles, as well as the most commonly used metadata standards in archival description.

Moving on, the focus shifted to accessing oral history content. To address this, the chapter briefly explained the principles of metadata, the OHASA oral history access workflow, and the formats of oral history content in general. The researcher then reviewed the literature on technologies used for metadata creation and access to oral history. This involved exploring software applications, finding aids, and databases. This study aimed to propose a metadata framework. Consequently, the discussion explored factors that influence the development or selection of the metadata schema to understand the framework best suited for oral history content.

The primary purpose of this review of the literature was to gain insight into current metadata practices related to oral history, both locally and internationally. The literature revealed that

challenges with metadata schemas and standards have a long history. However, the metadata schema has evolved since then and will continue to evolve. It was evident that efforts were being made in various countries to find ways to uniquely describe oral history for access and discoverability. It was also clear that no single schema or standard is compatible with the metadata needs of oral history content. Challenges stemmed from factors such as the language of metadata and that of the oral history content, formats of oral history content, types, and elements required for metadata of oral history content, as well as ethical issues related to access.

The literature acknowledges this lack of standards and the associated issues of interoperability, quality, use, and access, but the literature focused particularly on oral history metadata and description standards and solutions to these issues is sparse. This study seeks to fill this gap and understand oral history metadata for use among oral history collections. The next chapter focuses on the research methodology applied in this study.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 INTRODUCTION

The preceding chapter examined the literature on metadata schema for oral history content. It also discussed the metadata schema and standards that were used, access to oral history content, and the technologies that were available for metadata creation and access to oral history content. The purpose of this study was to investigate the metadata schema of oral history content in South Africa, in order to propose a framework to ensure its access and retrievability. To this end, this chapter focuses on explaining and justifying the research methodology for this study. Research methodology, according to Wellington (2015), is a notion that takes into account and clarifies the reasoning behind various research methods and techniques. Leedy and Ormrod (2015) refer to research methodology as the procedures used to identify, analyse and select information on a subject. A research methodology section gives the reader the opportunity to assess dependability and validity of the study. This is also important for ensuring the accuracy of the research conclusions that will be arrived at in this study, while maintaining that data were obtained through reliable procedures.

The methodology section addresses two main questions: 'How was the information collected or generated?' and 'How was the information examined?' This chapter provides an explanation of the research paradigm, the research approach and methods, quality criteria, data analysis, and ethical considerations as applied in this study. As reflected in Figure 3.1, this chapter covers the philosophical paradigm, research approach, research design, population and sampling, data collection techniques, data analysis, research rigour, ethical considerations, and evaluation of research methodology.

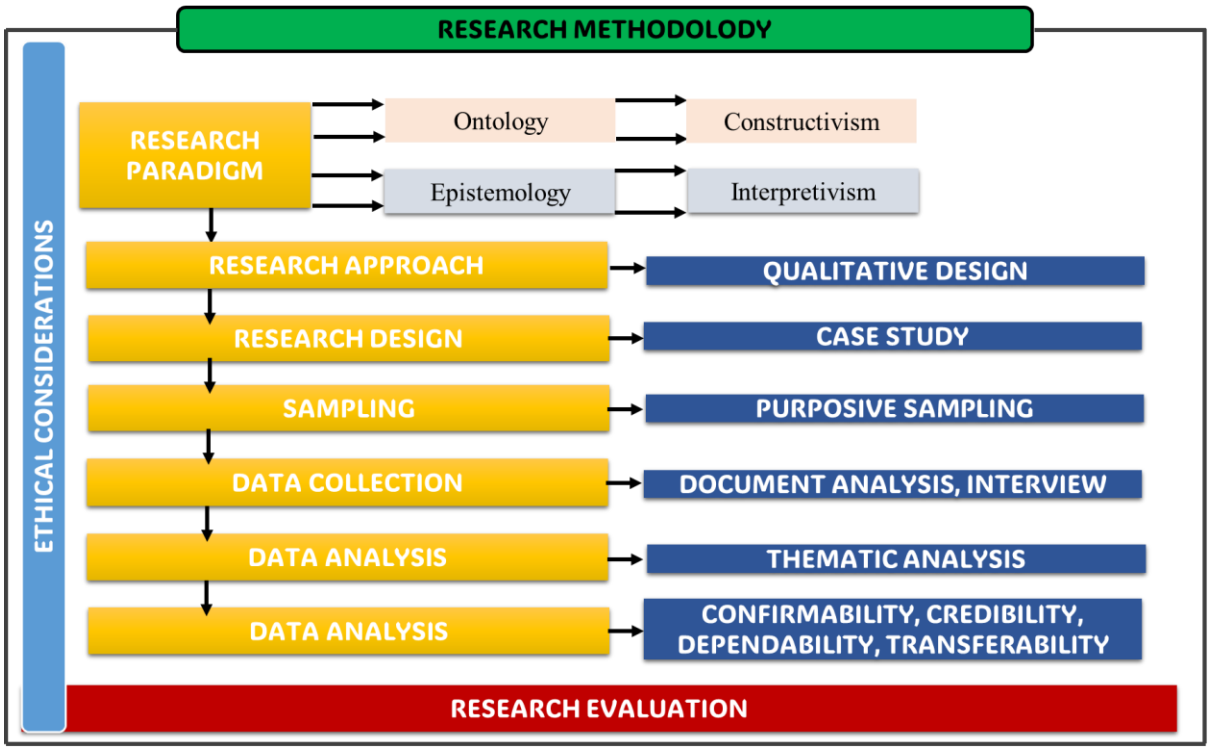


Figure 3.1 Research map.

The next section discusses research paradigm for this study.

3.2 RESEARCH PARADIGM

The research paradigm constitutes a crucial element within the research methodology that guides the effective and appropriate collection of data. Creswell and Poth (2018:35) define a paradigm as the “the whole constellation of beliefs, attitudes, and procedures held by members of a specific community." This paradigm is important, “as it provides the philosophical foundation for a research, and shapes the approaches used by different academic disciplines, in their research projects”, (Denzin & Lincoln 2018:164). Research paradigms essentially

encompass the philosophies of science and mould the fundamental aspects of scientific practice. In research, paradigm has four elements and/or positions, namely, epistemology, ontology, methodology, and axiology (Du Plooy-Cilliers, Davis & Bezuidenhout 2014; Kivunja & Kuyini 2017). These components make up the fundamental presumptions, convictions, standards, and values that each paradigm is based on, so it is crucial to have a solid understanding of them (Kivunja & Kuyini 2017). Simply put, paradigm includes ontology, which shapes the perspective on reality; epistemology, which dictates the understanding of knowledge; axiology, which determines the role and values attributed to the research process; methodology, which outlines the processes entailed in conducting research; and rigour, which entails the benchmarks employed to validate research quality within the paradigm.

Epistemology and ontology, which are fundamental components of any research, require a clear elucidation at the beginning of the methodology. These philosophical underpinnings significantly influence the theoretical approach used to address the research questions (Bryman 2016).Thus, the beginning of this section on research paradigms requires a comprehensive explanation of both epistemology and ontology. The section after that delves into a discussion of epistemology, followed by an exploration of ontology.

3.2.1 Epistemology

Epistemology refers to “the branch of philosophy that focuses on the nature of knowledge and the process used to obtain and verify knowledge acquired” (Gall, Gall & Borg 2013:16). According to Creswell and Plano Clark (2018), epistemology provides a conceptual framework for determining what types of knowledge are plausible and how to confirm their validity. Saunders, Lewis and Thornhill (2019) agree that epistemology is a way of viewing the universe and deriving meaning from it, and thus includes a specific understanding of what that knowledge entails. In earlier years, Cohen, Manion and Morrison (2018) emphasised that epistemology is mainly focused on “the nature and forms [of knowledge], how it may be acquired, and how it can be communicated to other human beings.” In order to understand "how things really are" and "how things really work," one must embrace an objective detachment or value if a single verifiable truth is assumed. Bryman (2016) states that epistemology pertains to the investigation of what is (or could be) viewed as adequate knowledge in a particular study. This suggests that epistemology has to do with how researchers investigate knowledge of a specific phenomenon in a study. The three dominant

epistemological philosophies that inform research are interpretivism, positivism and pragmatism (Ngulube 2015; Bertram & Christiansen, 2014; Du Plooy-Cilliers, Davis & Bezuidenhout 2014).

3.2.1.1 Positivism

Positivism is an 'epistemological position that emphasises that knowledge of a social reality is based on what can be observed and documented rather than subjective understandings' (Matthews & Ross 2010:478). This research paradigm is mostly employed to assess and confirm ideas and/or hypothesis as it aims to examine, validate and forecast law-like patterns of behaviour. Positivists believe that there is one reality in the world of which we are all a part (Rehman & Khalid 2016). Positivists hold that since the world is real and "out there," relationship between objects/things can be quantified. Bertram and Christiansen (2014:23) state that positivist researchers often concentrate on the objectivity of the research process to avoid being biased by not allowing their personal values and beliefs to interfere with the research.

Positive researchers use large sample sizes and experimental methods to produce generalisable facts (Bertram & Christiansen 2014). They use quantitative methodologies like survey, longitudinal, cross-sectional, and experimental methods, Bertram and Christiansen (2014). According to Johnson and Christensen (2014), the strength of positivism lies in its ability to generalise theory and collect data for the same issue in different social contexts.

Within the positivist paradigm, researchers posits that there is an external reality and a sense of order, as such they often employ quantitative and statistical data analysis in their research. Rehman and Khalid (2016) argue that research strategies generally fall within a continuum of positivism, including quantitative, scientific experimentation, and traditional approaches. Typically, positivists propose a hypothesis that can be proved or disproved by statistical data analysis. Positivism focuses on determining the relationships between variables, rather than the reason behind them.

Positivism is a research paradigm that 'explains and predicts what happens in the social world Positivism is a "research paradigm that looks for regularities and causal links between its constituent pieces in order to explain and anticipate what happens in the social environment" (Bonache & Festing 2020:104). According to Ngulube (2020), positivists believe that there is

an objective reality that is governed by natural causes. In positivism, the researcher maintains independence from the subjects of his study and strives to prevent bias by preventing his personal values and/or views from influencing the findings (Bertram & Christensen 2014:23). Most people link positivism with quantitative approaches (Bonache & Festing 2020). This study differed from others in that it did not focus on high sample numbers, instead focusing on causal links between variables to determine the impact of intervention (Bertram & Christiansen, 2014). Consequently, positivism is associated with quantitative methods (Bonache & Festing 2020). Typically, the main focus is on large sample sizes, which an emphasis on causal linkages between variables to determine the impact of intervention (Bertram & Christiansen, 2014), which was not the case in this study. Therefore, positivism was not applicable to this study on the basis that it was not aimed to determine the relationship between metadata schema and its variables, but rather intends to understand the reasons behind metadata schema and access to oral history content.

3.2.1.2 Interpretivism

Interpretivism is an epistemological put emphasis on individuals' subjective interpretations and understandings of social phenomena, as well as their actions (Matthews & Ross 2010:476). As indicated by Bibi, Khan and Shabir (2022), the interpretivists believe that reality is specific to each person and to the ways in which people construct, experience, and/or interpret their environment in their own particular circumstances and life experiences. The goal of social research, according to interpretivists, is to comprehend the meaning that underlies human behaviour. They maintain that there are a number of local, historical, local, particular, and generalisable realities and truths regarding the social world rather than a single reality or truth (Bibi, Khan & Shabir 2022). This suggests that events and circumstances can be interpreted in a variety of ways. Thus, it is acknowledged that research findings are produced by the interpretation of data rather than existing "out there" for the researcher to find or discover (Pervin & Mokhtar 2022).

In this regard, researchers make interpretations to better understand attitudes, behaviour, beliefs, human agency, and perceptions. This may in turn affect the choice of research methods. It seems to make sense that meaning can be understood through communication and interaction between researcher and participants. It should be recognised as a result that there would be subjectivity in the interactions between the participants and the researcher. Maree (2017)

explains that interpretivism places more emphasis on the subjective interpretations and understandings of social phenomena and individual actions by participants. This may thus suggest that interpretivism emphasises significance and meaning that individuals assign to their experiences.

Interpretivism rejects the idea of a single, verifiable reality, focusing on multiple socially constructed realities, and thus acknowledges that truth and reality are created, not discovered (Rehman & Khalid 2016). As reality is continually mediated by our senses, it is impossible to know. Observers cannot directly access the external reality without using their worldviews, concepts, backgrounds, and other factors to taint it. According to Flick (2018), it is believed that perception is an active, constructive process rather that a passive, receptive process of representation, whereby people engage with one another and with society to provide names and meaning to various social phenomena.

In the case of several well-reasoned interpretations of the same phenomenon, there is no one "correct" interpretation that is favoured over the others; rather, the existence of multiple bodies of knowledge is acknowledged with the understanding that different researchers approach the same problem from different perspectives. Interpretive research seeks to understand people's interpretations of the social phenomena they engage with, rather than seeking to discover knowledge and truth that is universal and context free (Maree 2017). Interpretive ontology focuses on understanding social phenomena through multiple perspectives, collecting qualitative data over extended periods (Al-Saadi 2014).

Researchers use inductive analysis to discover patterns in data, forming theories based on broad themes, as seen in ethnography and case studies. This is the opposite of the deductive approach, which involves identifying themes and patterns and themes before data collection, followed by searching for instances of these themes in words, statements and events.

Interpretivists use an inductive approach, viewing theory as derived from data collection, rather than statistical analysis (Khaldi 2017).They primarily use "verbal, non-statistical data, often recorded for authenticity, and use qualitative methods" (Gall, Gall & Borg 2003:21). One may thus conclude that interpretive researchers use methods that provide qualitative data, while numerical data may be used, it is not a primary source of information.

3.2.1.3 Pragmatism

Pragmatism as an epistemological position that can be used in mixed methods research (Ngulube 2015). Pragmatics supports the idea that a philosophy can only be valid if it functions in an acceptable way (Cassell, Cunliffe & Grandy 2018). It also embraces the notion that impractical ideologies should be rejected and that philosophy's significance can be discovered in the real world consequences of adopting it (Creswell 2014). Because pragmatism offers an alternative viewpoint to positivism, post-positivism, and constructivism and focuses on the problem to be studied and the results of the investigation, it is linked to mixed methodologies research (Teddlie & Tashakkori 2009; Creswell & Plano Clark 2018).

As previously mentioned, the pragmatic paradigm is a set of ideals that emerged as sole paradigm response to the discourse around the "paradigm wars" and the emergence of mixed methods and mixed models methods (Creswell & Creswell 2018). It is built in a variety of ways around the debunking of the forced choice between constructivism and post-positivism (Johnson & Christensen 2014). According to Jason and Glenwick (2016), the pragmatic paradigm has what they consider to be a natural plea regard as a natural plea and permission to study the relevant topic, includes acceptable approaches, and makes beneficial use of the findings in line with the researcher's value system (Creswell 2014). For this purpose, the researcher contends that the pragmatic paradigm is consistent with the combined quantitative and qualitative approaches carried out within the parameters of a specific research. This was not the case in this study, based on the view of Ngulube (2015) that the pragmatism approach can best be applied in mixed methods research. This study is a qualitative in nature.

The next section focuses on ontology as another philosophical perspective.

3.2.2 Ontology

Ontology refers to "the nature of reality, encompassing assertions about what exists (Creswell & Poth 2016:20). Researchers question existence of reality and knowledge through ontological questions, focusing on our beliefs about reality and the study of being, aiming to understand what exists in the world (Silverman 2020). These ontological questions are known as the "study of being" and they focus on "what actually exists in the world that humans can acquire knowledge about" (Moon & Blackman 2014). One may state that ontology assists researchers understand the nature and existence of phenomenon in their research. This is in line with

Creswell and Poth (2016)'s assertion that ontology is concerned with "what type of universe we are scrutinising, with the status of actuality, and with the structure of realism as such." Simply put, ontology addresses the question, "What is reality?" That is, does one's research include a single reality? A belief system that demonstrates "the way an individual perceives what constitutes a fact", (Don-Solomon & Eke 2018). Don-Solomon and Eke (2018) further indicate the two key positions in ontological philosophy that allow for the science of being or actuality of the nature of a social phenomenon under investigation as objectivism and subjectivism (constructivism) philosophy of research.

3.2.2.1 Objectivism

Objectivism as a ontological position indicates that " social and business entities exist in reality external to social or business actors concerned with their existence" (Kivunja & Kuyini 2017). This ontological position asserts that social phenomena and their meanings have a reality that is independent of social actors' (Bell, Bryman & Harley 2022). Objectivists reject the appearance-reality by holding that everything can be observed and understood as such without the need for social actors to mediate or meddle (Lincoln, Lynham & Guba 2011). According to ontology, regardless of the viewpoint or beliefs of the researcher, there is only one objective reality for any given occurrence or circumstance (Kivunja & Kuyini 2017). As a result, objectivists approach research with control and structure by selecting a precise and clear research topic, formulating relevant hypotheses, and using an acceptable and suitable research methodology suitable research methodology (Creswell 2013).

Since quantitative methods are objective and produce results that are repeatable and generalisable, they are typically used in research to achieve objectivism (Neuman 2011). Rehman and Khalid (2016) contend that the expectation is unrealistic since the five senses do not and cannot provide information that is independent of the concepts used to categorise it. Rehman and Khalid further indicate that objectivity is only achievable in situations where there is no mediating factor that skews or alters the observation. Moreover, concepts are what we naturally employ to characterise observations, and concepts always influence the result (Creswell 2014). Hence, this ontological position was not applicable to this study, since the focus of this study was to explore the metadata and access to oral history phenomenon, and as such there was a likelihood that the experiences and senses oof the participants would affect their individual observation and understanding of their existing realities.

3.2.2.2 Constructivism

Constructivists hold that there are multiple realities rather than a single reality or truth. Constructivism is about "the awareness that reality is a creation of human intelligence that interacts with experience in the real world' (Elkind 2004). Rehman and Khalid (2016) agree that constructivism accepts reality as a construct of the human mind; therefore, reality is perceived to be subjective. Rehman and Khalid (2016) add that constructivism aims to provide the answer to "why". Constructivism is recognised as a significant philosophical perspective that maintains a mutual relationship and dependence between the researcher and the social reality being studied (Moses & Knutsen 2019). Therefore, it can be concluded that constructivists focus on understanding and interpreting the meaning attached to an action. Table 3.1 summarises the ontology and epistemology as explained in this section.

Table 3.1 : Epistemological and ontological positions

***Epistemological stance: Nature of knowledge and how is* acquired**	***Ontology: Nature of reality and its existence***	
	Positivism	Closely linked to objectivism: Reality exists external to the researcher and must be investigated through a rigorous process of scientific inquiry, and can be quantified and generalisable (Matthews & Ross 2010; Bertram and Christiansen 2014; Rehman & Khalid 2016; Ngulube 2020; Kumatongo & Muzata 2021)
	Interpretivism	Linked to constructivism: Reality is specific to each person and to the ways in which people construct and/or interpret their environment, focusing on their meaning and interpretations, often use verbal non statistical data and qualitative methods (Matthews & Ross 2010; Al-Saadi 2014; Maree 2017; Zahle 2021; Bibi, Khan & Shabir 2022)
	Pragmatism	Offers an alternative viewpoint to objectivism and constructivism: External, multiple, view chosen to study a problem and the results of the investigation, is linked to mixed methodologies research (Teddlie & Tashakkori 2009; Ngulube 2015; Creswell & Plano Clark 2018; Creswell & Creswell, 2018)

The researcher should choose methods and methodologies that they find most suitable for investigating the phenomena they set out to investigate. Hence, the next section focuses on the rationale behind the choice of the research paradigm in this study.

3.2.3 Choice of the paradigm

This study adopted the interpretivism paradigm as a philosophical framework. Interpretivism is the study of the beliefs, motives, and reasoning of people in social situations as well as data that may be gathered about phenomenon (Bryman, Bell, Hirschsohn, Dos Santos, Du Toit, Masenge, Van Ardt & Wagner 2014). It is "commonly known that interpretivism involves employing qualitative methods" (Zahle 2021:202). Although the interpretivist researcher assumes that the complexity, plurality, and unpredictable nature of what is perceived as reality make it impossible to construct a fixed study design, they do approach the field with prior knowledge of the research setting Interpretivism is the study of the beliefs, motives, and reasoning of people in social situations as well as the data that may be gathered about a phenomenon (Bryman et al 2014). It is well knowledge that interpretivism involves using qualitative methodologies (Zahle 2021:202). Although the interpretivist researcher assumes that the complexity, plurality, and unpredictability character of what is perceived as reality make it impossible to construct a stable study design, they do approach the field with prior knowledge of the research setting (Bryman et al. 2014). Throughout the study, the researcher keeps an open mind and allows new information to emerge with the participants' assistance. Interpretivism is the study of the beliefs, motives, and reasoning of people in social situations as well as the data that may be gathered about a phenomenon (Bryman et al. 2014). It's well knowledge that interpretivism involves using qualitative methodologies (Zahle 2021:202). Although the interpretivist researcher assumes that the complexity, plurality, and unpredictability character of what is perceived as reality make it impossible to construct a stable study design, they do approach the field with prior knowledge of the research setting (Bryman et al. 2014). Throughout the study, the researcher keeps an open mind and allows new information to emerge with the participants' assistance. Interpretivism is the study of the beliefs, motives, and reasoning of people in social situations as well as the data that may be gathered about a phenomenon (Bryman et al. 2014). It's well knowledge that interpretivism involves using qualitative methodologies (Zahle 2021:202). Although the interpretivist researcher assumes that the complexity, plurality, and unpredictability character of what is perceived as reality make it impossible to construct a stable study design, they do approach the field with prior knowledge of the research setting (Bryman et al. 2014). Throughout the study, the researcher kept an open mind to new information and allowed it to grow with the assistance

of the participants(Bryman et al. 2014), despite the researcher's prior knowledge of the phenomenon under study.

The perspective of interpretivists includes constructivists (Creswell & Creswell, 2018). "Constructivism is a research philosophy that enables the creation of new theories or models and concretes them through contextual interpretations" (Rehman & Khalid 2016:55). This suggests that constructivists believe in the reconstruction and creation of new realities (Moses & Knutsen 2019). Constructivism is applicable to this study since the aim is to develop a framework. In this way, existing standards will be reconstructed, enabling a new framework through contextual interpretations of qualitative data collected from the NFVSA.

The interpretivism paradigm guided the adoption of qualitative methods in this study, as its emphasis is on the analysis of metadata schemas for oral history content in South Africa. In this study, metadata schemas were analysed to understand how they facilitate access to oral history content, thus generating a new framework; hence, this study adopted inductive reasoning within the constructivism/interpretivism philosophical framework of the qualitative approach. Limited previous studies on metadata access to oral history and interest in this research in understanding emerging aspects of access to oral history content through appropriate metadata and related processes in archival institutions suggested that an inductive approach was relevant to this research (Saunders, Lewis & Thornhill 2019). This aligns with collecting qualitative data, learning about the interpretations humans make of situations, and being less concerned about generalisation.

The interpretive paradigm is also relevant, as it was linked with the constructivist nature of this qualitative study. The researcher interacted with NFVSA staff to understand their experiences with regard to the metadata schema and how they interpret these experiences. This assisted in describing and analysing metadata standards as well as technologies used for metadata creation, thus allowing access to oral history content. In addition to gathering information, the intention of the researcher was to listen to the archival staff members discuss their methods and views on the practices and current and future challenges they perceive when providing access to oral history content (Saunders, Lewis & Thornhill 2019).

For this reason, constructivists frequently employ qualitative research methods that emphasises providing other viewpoints, such as case studies and interviews. Therefore, it was found to be relevant because the current study sought to gain an in-depth analysis of the metadata schema

used for oral history in the NFVSA rather than just the relationship between the metadata schema and oral history content. The beliefs held by researchers often influence the choice to embrace a quantitative, qualitative, or mixed methods approach in their research. The next section focuses on the research approach adopted in this study.

Once a research philosophy has been identified, an appropriate methodology can be chosen. In the research process, it is essential that the researcher specifically define the objectives, goals, and issues the study seeks to address. This implies identifying a research-worthy problem and defining achievable objectives to guide the study. According to Bryman (2016), after constructing the research question, the subsequent step involves selecting philosophical perspectives to provide a theoretical framework within which the study will be situated. Epitomised by epistemology and ontology, research paradigms are characterised by these philosophical assumptions (Ebohon, Ajayi & Ganiyu 2021).

3.3 RESEARCH APPROACH

Creswell (2014:3) describes a research approach as a "plan and procedure for any study that covers the steps from general assumptions to precise procedures of data collection, data analysis, and interpretation." Three major categories of research approaches, namely, quantitative, qualitative, and mixed methods, are the foundation of scientific inquiry (Hammond & Wellington 2013). Their differences stem from the way the researcher selects, analyses, and interprets the collected data (Du Plooy-Cilliers, Davis & Bezuidenhout 2014:14). The method of choice should be predetermined but be according to the intended outcome, the nature of the study, and the type of data required (Silverman 2022:18). Therefore, it is necessary to first discuss the three categories of research approaches before explaining the choice thereof. The section begins with the quantitative approach, then discusses the qualitative approach, and lastly the mixed methods approach.

3.3.1 Quantitative approach

Quantitative research is a descriptive method that uses formal measures of beliefs, attitudes, intentions, and behaviour, such as questionnaires and systematic observation, to develop knowledge through positivist hypotheses, experiments, surveys, and statistical analysis (Neuman 2011). According to Creswell (2014:4), a quantitative approach uses to test objective theories by examining the relationship between variables. Creswell (2014) further indicates

that quantitative methods place emphasis on quantification in data collection and analysis, which is expressed in numbers, percentages, and tables (Babbie 2013). Research methods in a quantitative study include experiments, surveys, document analysis, and existing statistics (Neuman 2011:22).

The primary draw of the quantitative approach is its historical significance as the oldest research methodology capable of explaining, forecasting, and describing a research phenomenon (Neuman 2011). One of the strengths of quantitative approach is its capacity to organise massive amounts of complex data into graphical representations or numerical summaries addressing the research questions in a given study (Ngulube 2009). The drawback of this approach is that the set protocols that researchers have to adhere to, there is a little flexibility but prescribed procedure to follow in quantitative research studies. Moreover, Remler and Van Ryzin (2011:16) claim that quantitative studies can reveal the link that exists between two variables, but do not offer as much to knowledge of what is responsible sponsible for the relationship as qualitative research methods do.

3.3.2 Qualitative approach

Qualitative research is referred to as "the study of the nature of phenomena", which includes "their different manifestations, quality, the context in which they appear, and/or the perspectives from which they can be perceive," but exclude "their frequency, range, and place in an objectively determined of cause and effect" (Busetto, Wick & Gumbinger 2020:1). Bryman (2016:392) concur that rather than focusing relationships, qualitative researchers are more interested in how people and/objects interpret occurrences, and what meaning they assign to phenomena, including the quality of the experience.

Several methods or research approaches are applied to conduct variety of scientific research. A research method is a data collection approach that may be used to any research design and makes use of tools including questionnaires, interviews, and observations (Bryman 2016). Both quantitative or qualitative research methods are possible. The kind of data that were gathered and the methods used for analysis to distinguish the two (Remler & Van Ryzin 2011). Considering their individual advantages and disadvantages, each approach should be considered as a choice rather than as a rival (Terre Blanche, Durrheim & Kelly 2006). Depending on the type of data needed and the nature of the investigation, one might choose between the two methodologies.

Qualitative research is a descriptive method that focuses on observing and describing everyday behaviours, rather than relying on numerical data (Silverman 2022). It aims to explore sociological elements and understand people's constructed meanings of their world. Qualitative researchers study phenomena in their natural settings, interpreting phenomena based on the meanings people bring to them (Stangor 2011; Maree 2017).

Qualitative research offers flexibility in exploring information through explanations, descriptions, and narratives. It allows respondents to provide their views on metadata schema of oral history content at the NFVSA. This method is particularly relevant to studying social relations and helps determine experiences and perceptions of employees and managers (Flick, 2018). In short, qualitative approach provide the researcher with the opportunity to analyse and describe phenomenon the way it is understood. However, qualitative research has disadvantages including subjective nature, time-consuming data gathering techniques, and potential researcher bias due to its lack of statistical analysis and neutrality (Stangor 2011). For example, qualitative data gathering tools such as interviews and document analysis are often associated with research bias.

3.3.3 Mixed methods approach

Mixed methods research employs both quantitative and qualitative approaches to investigate a research problem (Ngulube 2020). It is a research approach in which the researcher collects and analyses data, mixes the results, and draws conclusions using both quantitative and qualitative approaches in a single study (Creswell & Creswell 2018). When a single approach is insufficient to fully address a complex research subject including several individuals, aspects, and situational characteristics, a mixed methods study is appropriate (Creswell & Creswell 2018; Ngulube 2020). Leedy and Ormrod (2015) state that to mixed method approaches as a mixed methods research strategy that combines enquiry approaches with theoretical perspectives that direct the process of data collection, data analysis, and interpretation as well as the integration of qualitative and quantitative data in a single or category study (Babbie 2013). The combination of these methods is used for the purpose of comprehension and justification of research data collected and analysed, add Creswell and Plano Clark (2018). In this way quantitative methods such as descriptive and statistics as well as qualitative techniques such as interviews are utilised are mixed to ensure validation and in-depth information in research (Saunders, Lewis & Thornhill 2019) while minimising

disadvantages of using one data collection technique (Bless, Higson-Smith & Sithole 2013; Leedy & Ormrod 2015). One of the most advantageous characteristics of conducting in mixed methods research is triangulation and utilisation of multiple ways (methods and data sources) to examine the phenomenon (Morgan 2019). However, this was not the case in this study. The next section discusses the research approach chosen for this study.

3.3.4 Choice of the approach

This study was carried out using a qualitative research approach. The qualitative approach is defined by Creswell and Poth (2016) as an understanding enquiry process based on a unique methodology approach to enquiry that investigates a social or human problem. This study provides a comprehensive picture, analyses metadata standards, provides extensive participant perspectives, and was conducted in a natural setting (NFVSA). According to Brynard et al. (2014), the qualitative approach is a research method that generates descriptive data, which are the participants' own written or spoken words about their experiences or perceptions. The qualitative approach was used in the study because the goal was to gain a clear in-depth understanding of the metadata classification schema of oral history content in South Africa. Therefore, this study used a qualitative approach to analyse metadata standards and used participants' own spoken words to understand their experiences with the metadata schema in a natural setting (NFVSA). This is because this study intends to obtain a clear and in-depth understanding of the metadata classification schema of oral history content in South Africa. Knowing what paradigm to be used is important as it 'identifies what questions are to be considered in research study and what processes are relevant for answering those questions (Du Plooy-Cilliers, Davis & Bezuidenhout 2014:9).

Qualitative research is conducted in an interpretive framework. Within the interpretivist paradigm, it is permissible to use any approach, including quantitative procedures (Creswell & Creswell 2014). This also informed the research design for this study.

3.4 RESEARCH DESIGN

The research design is a complete plan for the entire research project that outlines how one will be formulating the hypotheses and question(s) to collecting data and finalising the analysis (Kumar 2011:396). Bertram and Christiansen (2017) also refer to the research design as a plan for how the research will collect the data needed to answer the research questions. A good

research design allows researchers to 'answer questions accurately, objectively, validly, and economically.' It thus necessary that the section that follows discusses the research designs.

3.4.1 Research designs in qualitative research

There are a variety of research design strategies to be considered within the three approaches, qualitative, quantitative, and mixed methods approach. These research designs include case study, ethnography, grounded theory, phenomenology, and narrative research (Creswell 2014; Bertram and Christiansen 2017).

3.4.1.1 Case study

A case study is characterised by a focus on interactions, processes and relationships; a focus on natural settings; an in-depth study of one setting; holism; a concern for the particular; and multiple methods of data collection", (Denscombe 2014:54-57). Similarly, Strydom and Bezuidenhout (2014:179) state that case study research "allows a deep exploration within a natural context, understands a phenomenon within a specific circumstance, and provides a full and thorough understanding of the specific lived experience of a participant." On the other hand, Hancock and Algozinne (2017) indicate that case studies have geographical parameters and boundaries allowing for their definition, can be defined by an individual and/or characteristics of the group in a particular context at a time, and can be shaped by organisational or institutional arrangements. This is supported by Creswell (2014), who describes case studies as set in geographical, institutional, organisational, temporal, and other contexts that allow limits and boundaries to be drawn around the case.

Research employing case study have benefited greatly from interpretive approaches (Yin 2016). This is because this design is aimed at investigating real world research issues and phenomena (Creswell (2014). One may thus conclude that in a case study design, researchers should consult a wide range of sources of information and supporting documentation. It is recommended to access a large number of evidence and information sources. Yin (2016) and Turhan (2019) concur that research problems in this approach frequently relate to people's lives, social phenomena, organisations and individuals. Creswell (2014) summarises the steps in case study as writing research objectives, developing subgoals, selecting analytic units/s,

elucidating the scenario, selecting participants, gathering information and determining links between objectives, evaluating and interpreting data collected and reporting the findings.

3.4.1.2 Ethnography

Ethnography refers to a method of identifying a group that shares a culture and examining how it evolves shared behavioural patterns over time (Creswell 2014). Creswell (2014) further indicates that ethnography enables researchers to observe participant behaviour while they are participating in events. Turhan (2019) adds that ethnography as a scientific approach that looks at how people interact with each other in the natural environments, thus researchers hope to build a closer relationship with the people under study, and their behaviours are explained by connections to their cultural setting (Turhan 2019).

While there are many different data sources used in ethnography, including watching what oc curs, hearing what is said, and asking questions about what is observed, the most crucial aspe ct is the emphasis on time as a dimension. Simply said, ethnography is sees as a "long term physical and/or virtual observation" of participant behaviour (Flick 2022: 467).

Grounded theory

Grounded theory design, also known as 'embedded theory development' and sub theory, is the examines the construction of social processes and structures, and a way of data collection techniques are methodically gathered for analysis (Creswell 2014). This suggests that in grounded theory, the researcher may develop a theory as the ultimate outcome of the study and present at the end or as a conclusion of the research project.
Grounded theory emphasises the importance 'reciprocal relationship between data and theory since it is necessary to give data the freedom to produce propositions in a dialectical way that allows for the application of preexisting theoretical frameworks while preventing any one framework from serving as a vessel for the facts', add Creswell (2014) and Turhan (2019). Thus, the researcher assumes that new phenomena may be found and that evidence for one or more unidentified theories may be found.

3.4.1.3 Phenomenology

Phenomenology as a "theoretical foundation for interpretive approaches", (Flick 2022), studies inquiries concerning personal encounters and/or experiences. The main aim is to define a

particular phenomenon, not generalising about it. Phenomenological inquiry clarifies experiences and the meanings attached to them, as well as to identify and investigate the existence of a phenomena (Turhan 2019; Silverman 2020; Flick 2022). This means that the research using phenomenological design, may concentrate on peoples' conscious experiences, everyday activities, and their social interactions. Turhan (2019) further indicates that one of the distinguishing features of phenomenology design is that every study participant has prior experience and information related to the research topic.

3.4.1.4 Narrative research

In social science research, a narrative inquiry or research is the study of experience understood narratively (Chung & Clandinin 2023). Litchman (2014:124) refers to narrative research as an "approach that relies on written or spoken words or visual representations of individuals." Creswell and Creswell (2018) add that a narrative research inquiry is employed across all social scientific social science disciplines, and it is used for providing the reader an insider's perspective of the environment. Creswell and Creswell (2018) mention some features of the narrative research summarised as follows:

- Narrative studies collect stories from people about their lived experiences.
- Narrative stories are collected through numerous ways such as interviews and documents.
- Narrative stories are analysed in varied ways, such as thematic analysis.

In summary, using narrative research is a qualitative approach , researchers examine people's stories about their experiences. Researchers should choose on a type of design within the chosen research approach. Sometimes it is not possible for research to best fit into one research design due to the nature of the study and the research topic at hand. As a result, the researcher adopted the research design of the case study design.

3.4.2 Choice of research design

The study adopted case study which is a research design that depict what it is like to be in that particular situation in order to capture the close-up reality, intricate detail, and "deep description" of the lived experiences, ideas, and emotions of the participants (Yin 2014).This study was aimed at conducting an in-depth analysis of metadata schema within its NFVSA context with a view to understanding the perspective and/or lived experience of the staff in

metadata creation and access to oral history content. Hence, this study was established within an organisational boundary, namely the NFVSA, and focused on the analysis of metadata standards and schemas used for metadata of oral history content from the NFVSA.

The researcher explored, understood, and presented the perspectives of the NFVSA staff and got close to them in their natural setting (Creswell 2014). The researcher interacted with NFVSA staff as necessary to generate data. The process of developing a research design begins with the location of the proposed study within a particular research paradigm. The case study design was also adopted because interpretivism commonly permeates the implementation of this research design (Litchman 2014; Harrison, Birks, Franklin & Mills 2017).

This study focused on the metadata schema of oral history reality as the lived experience of the NFVSA staff, with the aim of authentically representing their case and, unearth facts that amplify the distinctive voice of people whose experiences and viewpoints are often neglected, suppressed or unknown (Strydom & Bezuidenhout 2014). This study demonstrates how a case study recognised the context within which the oral history content metadata creation process unfolded, while the embedded narrative enquiry served to clarify the metadata schema and standards used at the NFVSA in South Africa. Therefore, it is important to explain the population and sample used to ensure and justify that the NFVSA represented the design of the case study of this study.

3.5 POPULATION AND SAMPLING

To achieve the main goal of any research, entities and/or units must be used to collect data. This section describes the population and proceeds to the sample and the sampling procedure in this study.

3.5.1 Population

Population refers to the group to which the research results can be applied or for which the findings can be generalised as clarified by the objectives of the study (Cohen, Manion and Morrison (2018). Creswell (2014) concur that population refers to a group and/or objects from which the researcher may draw conclusions. As a result, the population in any research should fit the characteristics relevant to the objectives under study, either directly or indirectly. Shukla (2020) defines a population as “that which consists of the sets of units that possess variable characteristics and for which the findings of research can be generalised and applied”. Wiid

and Diggines (2013:186) also define a population as "the total group of people or entities from whom information is required". According to Du Plooy-Cilliers, Davis & Bezuidenhout (2014), "one must return to the research problem and research question to determine what or who would best assist to find the answers needed to describe the population for the research, know whom to go to in order to find these answers, and ascertain what they have in common (what characteristic they all share) in order to define the population in a research proposal or report." According to these definitions, the population as a whole should share some characteristic relevant to the research question, and, most importantly, the population should be able to provide information and/or answers to the research problem and research objectives of the study.

The NFVSA staff formed the study population. However, the population of this study consisted of 16 staff members, eleven of whom were permanent and three of whom were temporary and one contract official, as well as one Head of National Film, Video, and Sound Archives. Two people worked in the legal deposit unit, two in the sound preservation unit; four in the film preservation unit; one in administration and registry; three in the outreach, client services and special projects; and three in the information system managements unit. The outreach unit oversaw oral history, with all other sections involved in various projects. Since oral history comes in a variety of formats, sound preservation is involved in the audio content of oral history.

These NFVSA staff were experts who could provide information (Creswell 2014) and had access to oral history content and were directly or indirectly involved with the creation of metadata for this oral history content (share characteristics). As such, they possessed characteristics relevant to the main purpose of this study, which was to investigate the metadata schema for oral history content in South Africa. The study population also included metadata records for the NFVSA oral history content. Metadata records were chosen to investigate the tools used to create metadata and to examine the standard for oral history content. The NFVSA hosts documents and metadata records as a legal deposit for oral history content and metadata from South Africa's ten public archives repositories, and it is mandated to be responsible for the country's oral history content. A well-defined population assists the researcher in selecting a sample of sufficient size to represent the entire population. The sample is crucial to the success of research and the reliability of the results. The following section focuses on sample and sample procedure to follow how the sample was chosen.

3.5.2 Sample and sample procedure

A sample is chosen when the entire population cannot be studied. This means that the units chosen from the population to represent all the characteristics of various types of populations, are referred to as samples (Shukla 2020). This implies that a sample is any part of the population that represents the population in all of its characteristics to be a complete representative thereof (Du Plooy-Cilliers, Davis & Bezuidenhout 2014). Some authors (Creswell 2014; Cohen, Manion and Morrison 2018) concur that a sample may be a subset of population represents all the different types of elements that make up a population. In this way a sample provides information about the entity from which it is taken. A good research sample is essential because (Shukla 2020; Kumar 2014; Creswell 2014; 2018), the research process becomes faster, less expensive, and precise due to fewer subjects in the sample; data collection is easier, and the researcher remains in control; and data analysis and interpretation are precise, and face-to-face contact becomes convenient.

A good sample should be chosen using specific methods. A sampling method is a method for selecting a sample. The research objective will determine the sampling method that can be selected by researchers from a variety of options (Shukla 2020). Sampling is defined as a "process of selecting a portion of the population to represent the entire population, while a sample is a subset of population elements for the purpose of investigation" (Polit & Beck 2010:307).The selection of such a sample makes the researcher's job easier and more precise. As a result, the sample size is critical in research (Shukla 2020). The researcher must carefully select the sample to collect accurate data for research; therefore, it is critical that the researcher use a specific sampling method with caution. For this reason, it is necessary to describe different types of sampling before justifying the sampling strategy used in this study. Different sampling methods are divided into two categories: probability sampling methods and nonprobability sampling methods.

A probability sampling method is one in which subjects have an equal or predetermined and certain probability of being selected in a sample, and subjects are chosen without bias or prejudice. In this way, units of the population have a fixed chance of being selected in a sample. In this method, subjects are chosen without regard for bias or prejudice. Due to its unique characteristics, it is considered the best method for selecting a sample (Shukla 2020). Non-probability sampling, on the other hand, is a sample selection method that increases the

likelihood of selecting biased samples. In most cases, such a sample does not accurately represent the entire population.

According to Kumar (2011:325), non-probability designs are used in qualitative research and are not limited by sample size. In qualitative research, sample is guided by and until data saturation is reached. Data saturation occurs when the accumulation of qualitative data reaches a point where no new analytical information emerges, leading to a sense of closure. Sample size in qualitative research is typically small due to data richness, participant variety, research question breadth, data collection method, and sampling strategy.

According to Thomas (2016), when the categories (or themes) are saturated, that is, when acquiring new data stops revealing new qualities or producing fresh insights, one should cease collecting data. Kumar (2011) adds that it is also critical to consider who the respondents would be and how they would be identified, that is, the determinants on which one would base a decision about the suitability of the respondents. The sample size is determined by the qualitative design used. Based on a review of numerous qualitative research studies, one can make an educated guess. According to Johnson & Christensen 2014), narrative consists of one or two people; phenomenology consists of three to ten people; grounded theory consists of 20-30 people; ethnography consists of one single culture-sharing group with numerous artefacts, interviews, and observations; and case studies consist of four to five units. Given that this study used a case study approach, a sample size of one to five people was sufficient to represent the population.

Polit and Beck (2010:307), on the other hand, state that "interpretivists tend to select a small sample of participant, ten or fewer." Polit and Beck (2010) add that the sample size is chosen based on the principle that individuals included in the study should be experienced in the phenomenon of study, enabling them to describe their lived that experience". The aim of qualitative research is to select participants or sites (or documents or visual material) that would be best aid the researcher in understanding the problem and the research question (Creswell & Poth 2016). According to Vasileiou, Barnett, Thorpe and Young (2018:148), qualitative research samples "tend to be small to support the depth of case-orientated analysis" and that qualitative samples are purposeful in nature, that is, "selected by virtue of their capacity to provide richly textured information relevant to the phenomenon under investigation."

3.5.3 Sample size

Since the total population was relatively small (16), the entire population was chosen for data collection because sampling processes are of no use when the sample size is fewer and/or less than 100 (Robinson 2014). Creswell (2015), underscore that in qualitative research, two to 10 cases are considered sufficient and valid for the study. Qualitative research is not aimed to select a random or unbiased sample, but to select one that can provide as much detailed, accurate, and complete information as possible. As a result, in selecting respondents, the researcher is primarily guided by their own judgment. All 16 NFVSA staff members who participated were well versed in metadata creation and oral history content or had both organisational experience and lived experience. This was also because there were fewer NFVSA staff members (NFVSA Annual Report). In addition to people, the researcher sampled other sources/objects, specifically metadata records. The metadata record was sampled to describe and analyse the metadata standards employed by the NFVSA. When the appropriate sample size is determined, the next step is data collection; therefore, the following section discusses data collection techniques.

3.6 DATA COLLECTION TECHNIQUES

In quantitative research, the main data collection tool is the questionnaire, whereas document analysis, interviews and observation are used qualitative research. (Moser & Korstjens 2018). Ngulube (2009) agree that qualitative research uses qualitative data from document analysis, in-depth interviews and observation, and ethnography. In this study, the researcher relied on interviews and document analysis (Appendix D). Although collecting data using two or more methods requires collecting and analysing more data, which requires more time, financial resources, effort, and technical expertise (Kumar 2014: 29), data instruments were required in this study to achieve the research objectives outlined in chapter one. The following section describes document analysis as one of the data collection techniques used in this study.

3.6.1 Document analysis

Researchers use document analysis as a qualitative research technique. Document analysis is important because it enables researchers to make sense of their resources and choose the best pieces for their studies. The choice of documents for analysis depends on the topic which contains data, ideas, information, and evidence to fulfil certain aims, and the nature of the topic

in relation to the research being proposed (Flick 2018). Documents contain qualitative data, which is non-numerical information organised into categories by researchers. Data organisation enables researchers to combine document analysis with other types of research. Document material is used by researchers because the information contained in it can be used for various purposes, such as (Laher, Fynn and Kramer 2019):

- Giving the research context
- Suggesting potential questions researchers may ask
- Recommending scenarios and/or situations researchers to be observed
- Providing additional valuable insights for researchers
- Providing researchers, a mechanism to monitor the progress of their studies
- Acting as a means of validating data from previous studies

Researchers frequently utilise the information gleaned from document analysis in a research procedure, which compares information from two or more sources in the same category to identify patterns and discrepancies. To create trustworthiness, this process is essential. In this study, document analysis was used to analyse information from guidelines, reports, and other relevant records (Creswell 2014). The document analysis checklist (Appendix D) was developed to collect data relating to the objectives of the study. To gain a better understanding of metadata schema and metadata standards, policies/ guidelines for metadata creation and metadata schema used for metadata creation of oral history content were examined. By analysing the metadata records, the researcher collected data relevant to understand access to oral history content and the fields that were used and/or not used to describe oral history materials. As indicated earlier, interviews were also used to collect data.

3.6.2 Interviews

This study adopted interviews as one method to collect data. According to Bless, Higson-Smith and Sithole (2013), an interview is a method by which the researcher initiates and organises direct special contact with the participants, with the assumption that the interviewee will answer questions provided by the interviewer during the data collection phase. An interview is a type of conversation that involves two or more people. Most of the time, one individual leads the interview by asking questions (Matthews & Ross 2010:219). Depending on the purpose of the interview, the interviewer might guide the interview in a highly structured or unstructured manner (Denzin & Lincoln 2018).

Cohen, Manion and Morrison Cohen (2018) summarise the research objectives of an interview as follows:

- It is the main means used to collect data for research purposes, including learning about people's thoughts, knowledge, preferences, values, and beliefs.
- It is employed as an explanatory tool to assist identifying factors and relationships, as well as to test and propose new ideas.
- It could be used in addition to other techniques.

Interviewing generates a greater amount of qualitative data. The advantages of conducting interviews encompass the researcher's capacity to scrutinise the environment, employ nonverbal cues, and inquire deeply. Face-to-face interviews have several drawbacks, including interviewer bias and question phrasing, look, and tone that might occasionally affect respondents' responses(Hammond, & Wellington 2013). Following the analysis of the document, semi-structured interviews were conducted with NFVSA staff. Given the circumstances at the time of data collection, interviews were conducted in person and/or virtually (via Microsoft Teams).

Information system managers and metadata archivists were regarded by the researcher as important people who may offer priceless insight and more thorough responses to study queries. Having access to archival materials, especially oral history content, was essential to their job. All respondents were asked the same questions in the same order and in the same format. This was in line with the suggestion made by Kenneavy, Harnois, Atkinson, and Korgen (2022) to ensure that interview subjects are given questions that are similar to each other, a structured interview guide should be used. This type of interview minimises volatility and flexibility while maximizing consistency (Punch 2014:146).

The researcher found information system managers and metadata archivists as key individuals who could provide invaluable insight and more detailed answers to research questions. Their work mostly entails having access to archival materials, including oral history content. This was consistent with Kenneavy, Harnois, Atkinson, and Korgen's (2022) recommendation that a guiding structured questionnaire should be used in interviews to ensure that participants are asked similar questions. Flexibility and variation are minimised in this type of interview, while standardisation is maximised (Punch 2014:146). The interview schedule was divided into four parts: Section A: Background information; Section B: Metadata schema; Section C: Metadata

standards; Section D: Access to oral history content; and E: Technologies used to access metadata and oral history content (See Appendix D). The interview schedule had 26 questions aligned to the objectives of the study. The schedule included also included questions on the background of the participant to provide a picture of people being interviewed. When deciding on an interview schedule, the researcher took into account the following benefits of the interview schedule as stated by (Kumar 2014:182):

- Because there was no restriction to the amount of space in the questionnaire or observation schedule, the interview helped to elicit more detailed information from the respondents.
- Data could be expanded upon. As a result, the interviewer was able to add information from questionnaires and observation to that which was gathered via responses.
- It is possible to explain questions. Because the interviewer could either repeat a question or phrase it in a way that the respondent understood, there was less chance that it would be misunderstood.
- The participant's likelihood of preparing to lie in his response was decreased by the interviewer-interviewee exchange.

In this study, when collecting data through interviews, the researcher encountered the following disadvantages:

- Both costly and time-consuming. The researcher in this study had to postpone the interviews because the oral history staff was consistently unavailable because the projects were being undertaken outside of the NFVSA (travel to different provinces). The researcher had to visit the location again and send an email.

3.6.2.1 Recording of interview data

To obtain the most accurate direct measurements, the researcher considered recording as a data collection technique (Walter 2019: 284). Recording and taking notes are two options used for more open-ended interviews. These various options were evaluated in view of the practical constraints, the respondents' approval, and the type of interview chosen. The researcher then decided to use the method of taking notes and recording. Although the interviewees agreed to being recorded, the researcher found it difficult to follow up with some of them because they spoke too fast. To address this flaw, the researcher had to ask the interviewees to slow down.

The researcher captured and transcribed the data to ensure that everything was written exactly as it appeared on the notes.

The interview procedure, which includes the following essential elements, was adhered to by the researcher: general interview information, introduction, interview content questions with probes, and closing instructions (Creswell 2016). The researcher planned to develop and use an interview protocol for capturing questions and answers during the qualitative interviews. To capture information, he recorded interviews on audio and collected handwritten notes. Researchers should take notes during interviews even if they are videotaped, just in case the recording device malfunctions. In qualitative research, transcription is frequently a "essential and basic step" (Flick 2018: 438). According to Singleton and Straits (2018), a translation converts spoken words into written language and is typically a partial record of an interaction.

Researchers need to schedule in advance for the transcription of the devices if audiotaping is utilised. It must to be ready ahead of time and used consistently to every interview. To avoid coming across as someone who is just reading the interview protocol, it is beneficial for the interviewer to commit the questions to memory.

3.6.3 Data collection procedure

Numerous qualitative studies include the collection of various data types and a substantial amount of fieldwork time by researchers to acquire information. There exist four fundamental categories of data collection processes in qualitative research (Creswell & Creswell 2014; 2018):

(a) Qualitative observation involves researchers taking field notes on individuals' behaviour and activities at a research site, recording unstructured or semi-structured activities. These observations are open-ended, allowing participants to freely provide their views through general questions.
(b) In qualitative interviews, the researcher conducts face-to-face interviews, telephone interviews, or focus group interviews with six to eight participants in each group. The interviews in this study involved a few unstructured and generally open questions that were intended to elicit views and opinions from NFVSA staff.

(c) The researcher can gather qualitative documents, including public and private ones, during the research process. In this study, the metadata records were used as qualitative documents.

(d) A final category of qualitative data comprises of qualitative audio-visual and digital materials (including social media materials). The data in this study also took the form of the main pages of the NARSSA website, where the NFVSA web page was located (Cresswell & Poth 2016).

To conduct investigations, researchers must follow standard data collection procedures. The following procedures were followed by the researcher. First, the researcher applied to the University of South Africa for an ethical clearance certificate (Appendix C). After receiving the certificate, the researcher sent an email to the NFVSA requesting permission to conduct research. The head of the organisation granted approval and assigned the researcher an official who was in charge of facilitating access to participants, coordinating appointments with participants, and gaining access to documents relevant to the study. The participant information sheet (Appendix A) was given to the respondents, but they were also discussed before signing informed consent (Appendix B). Details of the ethical issues addressed in this study are covered in paragraph 3.9, Ethical considerations.

As part of the data collection preparation, the researcher visited the site for a general conversation about the organisation. The meeting provided the researcher with a better understanding of the topic under investigation and allowed him to refine some of the questions. At this point, some individuals outside the organisation were mentioned. Initially, the researcher did not contact these individuals, but after several interviews with the respondents, he saw the need to also schedule interviews with these individuals. One of the objectives was to analyse access to oral history content and the technologies used for metadata creation and access to oral history content, and NARSSA had an information management system. This is because, as stated earlier in this study, the NFVSA is a sub-directorate of NARSSA and is mandated to collect oral and audio-visual collections, including oral history. There was a requirement to obtain information from the information system manager at the parent organisation. Because a face-to-face interview was not possible due to Covid-19 regulations, an online interview was conducted, recorded, and transcribed using Microsoft Teams. The interviews were conducted face to face and through Microsoft Teams at a convenient time for the participants and lasted no more than one and a half hour. Participants received regular

breaks, but none requested that the interviews be paused. As there was no strict and tight pattern of asking questions, the researcher allowed each interview to be guided by its flow, with each question being followed by probing questions; however, the interview schedule (see Appendix D) guided the process. The interviewer attempted to engage in active listening throughout the interviews, being attentive, responsive, and curious (Silverman 2022).

Moreover, despite the fact that the meeting was recorded, notes were also taken during interviews.

The researcher also visited the site to examine the documents. Because some documents, such as NARSSA policies and legislation, were available online, the researcher accessed them there. They also observed oral history content formats available at the NFVSA. It was critical to request and schedule visits around the respondents' schedules in order to respect the site and disrupt daily activities as little as possible (Creswell & Creswell 2018). Creswell and Creswell (2018) further state that researchers must ensure that research sites are not disturbed after a study and must show respect of the site. This means that the researcher must be conscious of their influence and try to minimise disruption to the physical setting (Kumar 2011). The current researcher made sure that the visits were timed so that they did not disrupt the respondents' activities. This was also done because the NFVSA was understaffed and had to deal with the backlog in general, and specifically, oral history metadata. It was also the time when oral history projects were in full swing, so most of the staff were absent and the organisation was run by a skeleton crew.

This study used a combination of document analysis and interview data collection methods. This integration was put in place to make up for one data collection tool's inadequacies relative to another. Punch brought up two broad considerations that the researcher kept in mind during the data collection process (2014:242). The investigator made certain that the participants were handled with professionalism and propriety, and that they were fully apprised of the context and goals of the study, of the confidentiality and anonymity policies, and of the uses and recipients of the information they submitted. This degree of confidence guaranteed that responders understood precisely what was expected of them and encouraged cooperation. Consequently, the quality of collected data was increased.

Face-to-face interviews were conducted with NFVSA staff members and online interviews were conducted with NARSSA members, to whom the majority of interviewees referred. As

the information system managers were placed at NARSSA, it was critical to conduct interviews with NARSSA personnel. Due to meeting commitments, the FVSA's head was not interviewed and left the organisation during the data collection process. A new head was appointed, but the efforts to secure the appointment were futile. For consistency, the interviews were recorded using various gadgets, as well as manually on notebooks. Managers were often unavailable for interviews due to the nature of their demanding work schedules. Because the interviews were postponed several times, the data were collected over a five-month period. During this time, the researcher communicated with the coordinators who helped to confirm the dates of rescheduled interviews.

3.7 DATA ANALYSIS

For this purpose, the researcher used thematic data analysis to identify themes. According to Braun and Clarke (2013), "thematic analysis is defined as a detailed and systematic examination of the contents of a particular body of material for the purpose of identifying patterns and themes." As stated by Polit and Beck (2010:477), "detecting patterns and regularities of themes as well as inconsistencies (done to analyse commonalities across instances)" is the task of a thematic analysis (Nowell, Norris, White & Moules 2017). To corroborate the findings, the researcher additionally provided sample response cases in research reports. Leedy and Ormrod (2013) identify steps taken by a qualitative data analyst. The researcher first identified the specific content to be studied and then defined the qualities to be studied. The researcher also divided the terms into smaller, more manageable segments that could be analysed separately. Finally, the researcher examined paying particular attention to the qualities defined in step two. This study employed a comprehensive research design based on document analysis of metadata records from the NFVSA oral history collections in South Africa. The analysis focused on specific description methods and metadata elements in use, as well as a close examination of existing records in each collection that present description and metadata.

In research, data analysis is influenced by the kind of research strategy selected for the process (Creswell 2009:218). According to Silverman (2020), there are several steps involved in data analysis that include arranging the case details logically, classifying the data, looking for patterns and underlying themes in the data, analysing the data for generalizations that emerge, and finally synthesising the findings.

Finding themes, or significant or intriguing patterns in the data, is the aim of a thematic analysis. These themes are then used to discuss the research or make a point. A strong theme analysis interprets and makes meaning of the data in addition to summarising it. A common error is to use the primary interview questions as themes (Clarke & Braun, 2013).

The researcher began by reading and re-listening to the transcripts, as is customary in thematic analysis (Siverman 2020). However, the researcher did not use a strict step-by-step approach to coding and thematising the data; rather, as the researcher read and re-read the transcripts, special attention was paid to the moments when the interviewees explained various aspects relating to metadata and access to oral history. This led to various themes being developed in accordance with the study's objectives.

3.8 RESEARCH RIGOUR

Quality criteria refer to the evaluation of the fit between the research methods and the research question, the target population, and the measurement of outcomes (Du Plooy-Cilliers, Davis & Bezuidenhout 2014:27). The qualitative method was used in this study, within which the researcher addressed credibility, confirmability, dependability , and transferability to ensure the quality of this study. Although commonly used in quantitative studies, reliability, validity, and objectivity were also addressed in this study because they were recognised as a need in qualitative research.

3.8.1 Reliability and validity

Reliability is concerned with the consistency of the measurements whereas validity is concerned with the precision of the study measurements (Rose & Johnson 2020). This section describes the reliability and validity of this qualitative study, beginning with explanation of reliability, followed by validity.

3.8.1.1 Reliability

Reliability in qualitative research is described as "the consistency and/or stability of replies to multiple coders of data sets." Using recording equipment and transcribing the digital files can improve the quality of detailed field notes, explains Creswell (2014). Yin (2016) asserts that qualitative researchers ought to record as many stages as they can in the techniques they use for their case studies. Additionally, he recommends building a comprehensive database and

case study protocol so that the processes can be repeated by others. Gibbs (2007),Creswell (2014) and Yin (2016) propose a number of qualitative reliability procedures.

- Verify the transcripts to ensure there are no glaring mistakes that were produced during transcription.
- Throughout the coding process, make sure that there is no drift in the definition of the codes or a change in their meaning. This can be achieved by drafting memos explaining the codes and their definitions and by continuously checking data with the codes (see the discussion on a qualitative codebook).
- Organise coder communication through documented meetings on a regular basis and by sharing the analysis when conducting team research.
- Multiple procedures for cross-checking codes should be developed by different researchers, ensuring consistent results. Intercoder agreement, or crosschecking, involves finding another person to cross-check codes for the same passages in the text, determining whether another coder would code it with the same or similar code. (Creswell 2016). An agreement may arise between coders regarding the codes used for the same text passages, rather than determining if another coder would use the same or similar code.

Qualitative reliability indicates that the researcher's approach is consistent across the study and researchers.

3.8.1.2 Validity

Validity, which is predicated on assessing if the conclusions are accurate from the viewpoint of the researcher, participant, or readers of an account, is one of the advantages of qualitative research (Lincoln, Lynham & Guba 2011). Validity in qualitative research, according to Creswell (2014), is the evaluation of the results' accuracy as best expressed by the researcher, participants, and readers. Cohen, Manion, and Morrison (2018) assert that evaluating validity involves addressing the nature of validity, defining validity, figuring out when one has reached a sufficient level of validity, managing validity in research jargon, and evaluating how validity impacts design, inferences, and conclusions (Laher 2016; Flick 2018). In qualitative literature, terms such as trustworthiness, authenticity, and credibility abound (Lincoln, Lynham & Guba

2011). Researchers can use one of the following validity strategies to ensure qualitative validity.

(a) *Triangulate* information from the sources to create a coherent case for themes (Creswell and Creswell 2018). The themes of this study were developed by the researcher using information gathered from interviews and document analysis. According to Creswell & Creswell (2018), a study's validity may be increased if its themes are derived through the synthesis of many data sources or participant viewpoints.

(b) Use member checking. Take the final report or certain descriptions or themes back to the participants and ask them if they believe the findings are correct. This technique is known as member checking. Using some of the analysed data, the researcher created themes that participants could verify for accuracy. At this point, the participants were given the chance to weigh in on the findings through a few follow-up interviews.

(c) *Reflexivity* is a fundamental feature of qualitative research. Good qualitative research includes comments from the researchers regarding how their background—including gender, culture, history, and socioeconomic origin—affects how they perceive the data.

(d) Provide contradictory or negative information that runs counter to the themes. To do this, the present researcher talked about and provided proof on the themes, both positive and negative. Evidence from the literature that supported or contradicted the themes improved this.

(e) *Spend a lot of time working in the field.* Visiting the location, engaging up a general chat, and earning the trust of the participants and the NFVSA allowed the researcher to acquire a thorough grasp of the phenomenon under investigation, even though they did not dedicate much time to it.

(f) Peer debriefing can help to improve the accuracy of the report. Finding someone who reviews and enquires about the study is part of this procedure to ensure that the story is meaningful to individuals outside the researcher. It was crucial for this study to have peers evaluate the data gathering to make sure that everyone's interpretation was the same. In order to confirm the reliability of the data collection instrument, the supervisor also verified the questions following several correction submissions.

Simply put, in qualitative research, reliability denotes consistency, whereas validity denotes trustworthiness. Researchers ought to address the general trustworthiness of qualitative research by offering a range of reliability and validity approaches that could enhance the credibility of analysis and findings portrayal, as opposed to dictating what reliability and/or validity ought to entail (Rose & Johnson 2020). Denzin and Lincoln (2018) regard reliability and validity as encompassing credibility, transferability, and conformability.

3.8.2 Credibility, transferability, and conformability

This section covers credibility, transferability, and conformability as applicable to this study.

3.8.2.1 Credibility

The credibility of the study is determined by the accuracy of results. By employing multiple methodologies and doing member checks, which enabled participants to confirm the accuracy of the interpretations in an interview with some of the staff members who helped create the metadata, the credibility of the study was increased.

3.8.2.2 Transferability

The transferability of a study lies with the practitioners and readers (Ary, Jacobs & Razavieh 2009). For this purpose, detailed information about where and how data were collected is provided.

3.8.2.3 Conformability

Lastly, *conformability* was improved by guarding against the researcher's personal values, bias, predispositions, and preferences that influence the conduct of the research by discussing with archival staff the metadata classification as it was developed. In addition, a complete audit trail was made available for other researchers to investigate the data and context used in the study.

3.9 ETHICAL CONSIDERATIONS

This section explains the ethical considerations that were followed in this study and how each was addressed throughout the study. Being ethical means "according to what is "deemed appropriate behaviour, particularly that of a particular profession or group" (Kumar 2011:217). Thus, we may deduce that for this study to be ethical, it was conducted in accordance with the principles that are considered correct in conducting research. Therefore, research ethics are

considered the core of any research. Johnson and Christensen (2014:192) concur that all research project must be developed and carried out with ethics in mind.

To understand research ethics, it is necessary to first explain what the concept of “ethics” means. According to Johnson and Christensen (2014), ethics refers to the guidelines and/or principles that assist one to uphold and comply to legal, professional, and social obligations to the study participants (Polit & Beck 2010; Thomas 2016). Thus, moral standards of behaviour that are often connected to morality and deal with issues of right and wrong are used to guide the decision-making behaviour of an individual or a group of individuals in research ethics (Agwor & Osho 2017). In this regard, ethical considerations “relate to systematised ways of considering what is right and wrong when conducting research" (Brynard et al. 2014:94). In their area of expertise, the researcher ought to make choices and follow rules of ethics. Jacobs (2020:53) states that “it is important to consider codes of ethics in the social sciences and humanities since the emphasis is on collecting data involving the use of humans or human documentation.” This study was conducted in the discipline of humanities, and collecting data involved interviews, the use of human personnel from the NFVSA, and human documentation related to metadata schema and creation. Therefore, it was relevant for this study to follow ethical considerations in this regard.

Creswell and Creswell (2018) point out that ethical issues should be addressed as they relate to different phases of enquiry. This means that a researcher should be aware that attention should be paid to ethical issues prior to conducting the study, during the beginning a study, during data collection and analysis, and in reporting, sharing and storing data (Creswell & Creswell 2018). Maree (2017:47) outlines ethical considerations when conducting research, including “permission to conduct the study, an informed consent form, anonymity and confidentiality, and avoidance of plagiarism.” Ethical considerations relevant to conducting this study are discussed in subsequent paragraphs.

3.9.1 Permission to conduct the study

Researchers must have their research plans reviewed and approved by their respective institutions. According to UNISA’s research policy, an ethics clearance application was submitted to seek permission to conduct this study. The ethics committee, among other things, requires research to assess the extent of potential risks to study participants in order to protect them from human rights violations and harm risk that may arise from the study (Creswell &

Creswell 2018). This study was reviewed and approved, and received ethical clearance (see Appendix C). After receiving ethical clearance, a permission to conduct research at the NFVSA was requested and approved. Before starting the study, it is important and beneficial to negotiate the publication plan so that people who contribute to the study are aware of what their contribution to their study means (Creswell & Creswell 2018). The researcher showed to the NFVSA that he had received an ethical clearance certificate from UNISA. After submitting the ethical clearance certificate to the NFVSA, the researcher was granted permission to collect data. Respondents who agreed to participate in the study were required to sign an informed consent form.

3.9.2 Informed consent

In order to collect data, the researcher and the participant must reach an agreement, which can be obtained through informed consent. In research, "coercion and/or pressure to participate" should not be used (Blaikie & Priest 2019: 55). According to Kumar (2011:244), informed consent implies that subjects are adequately informed about the type of information the researcher seeks from them, why the information is sought, what purpose it will serve, and how they are expected to participate in the study. Simply put, the participant should know fully and accurately about the nature of research. According to Creswell (2014) emphasise that participants should be informed about their participation in a research study. They should give their consent as they would have understood what would be required during their participation, as well as how their identities would be protected, and how the results will be used. Preferably, state this information in writing, have participants sign their consent, and keep these documents on record (Pickard 2013)

Pickard (2013:89) defines informed consent as "part of an agreement between the researcher and the participant that creates a mutual understanding that remains constant throughout the research." Participants were provided with an information sheet(Appendix A) about the purpose of the study reasons for participation, data collection tools, potential benefits, confidentiality, data security, and informed consent. They signed the consent form voluntarily, ensuring they understood the research and had the ability to consent or decline(Blaikie & Priest 2019; Pickard 2013). The consent form included contact information for both the researcher and supervisor, allowing participants to contact them if they had ethical concerns. Informed consent protects participants' right to autonomy(Litchman 2014).

According to Fouka and Mantzorou (2011:4), autonomy is the "ability to act on one's own initiative in accordance with a personal plan." This study promotes self-determination, protecting individual rights and personal liberty, thus participants were informed about the research process and given time to express concerns. Participants agreed to participate in a study after a thorough explanation of the research process, including an information sheet and permission to record interviews. They were given ample time to express concerns and questions, ensuring anonymity, confidentiality, and privacy during data collection (UNISA ethics policy).

3.9.3 Anonymity and confidentiality

Anonymity and confidentiality are crucial principles in research ethics. Anonymity is traditionally a "means of removing the contributor's name in order to protect them from harm mostly in qualitative studies which usually presents narratives or views of participants" (Dube, Mhlongo & Ngulube 2014:203). Saunders, Kitzinger and Kitzinger (2015) add that anonymity refers to 'keeping participants' identities secret and ensuring that a person will never be traceable from the data presented about them. In this way, the participants "identity is protected in such a way that they cannot be linked to personal responses" (Fouka & Mantzorou 2011:6). Anonymity occurs when a researcher cannot link a participant with their information, preventing them from being identified afterward (Polit & Beck 2014).

In social research, confidentiality and anonymity are closely related to each other. According to Polit and Beck (2010), anonymity is a means of securing confidentiality in research and to participants to their data. Wiles (2013:41) also states that "in social research, anonymity is the vehicle by which confidentiality is operationalised."

In research ethics, confidentiality entails an "obligation on the part of the researcher to ensure that any use of information obtained from or shared by human subjects respects the dignity and autonomy of the participant and does not violate the interests of individuals or communities" (Saunders, Kitzinger and Kitzinger 2015:617). This means that confidentiality in research ethics requires researchers to respect the dignity and autonomy of participants and avoid violating their interests. Most participants provide sensitive information, and it is their responsibility to treat it confidentially(Bos 2020).However, not all information should be kept confidential or anonymous (Lune and Berg 2017). Confidentiality can be achieved by not

sharing certain parts of the data or limiting the use of specific information if requested by the participant (Wiles 2013).

The study ensured confidentiality by restricting access to raw data, using pseudonyms, and distorting identifying details. It did not record participant names during the interview process, and if participant identity was known, data was de-identified and kept confidential. Access restrictions were also in place, and records were secured using passwords or encrypted files for internet transmission. This approach ensures no participant could be identified after the study. Confidentiality may also include restrictions to accessing data (Macleod & Mnyaka 2018). In a situation where the data had to be sent via the internet, records were secured by means of passwords and/or encrypted files.

Confidentiality is crucial in research, ensuring participants retain control over their privacy. Researchers avoid personal or intimate questions, conduct private interviews, and respect participants' dignity. They help participants anonymise interviews, ensuring easy identification and maintaining internal anonymity (Bos 2020), ensuring the privacy of participants.

Privacy is a principle of respect, allowing individuals to control their information(Pickard 2013:88). Invasion of privacy occurs when private information is shared without consent (Ngulube 2020). Researchers maintain privacy by introducing themselves, revealing study procedures, and not asking about intimate matters. They interview respondents privately, treating them with respect, and adhering to scientific integrity to ensure effective data collection. Considering the nature and autonomy of the organisation, it was easy to identify the interviewees. Hence, the researcher helped the participants to anonymise interviews, especially with respect to internal anonymity (Surmiak 2018). The researcher also followed scientific integrity to ensure that the data collected from the participants were effective and useful.

3.9.4 Scientific integrity

Researchers need to apply the principles of the scientific integrity of knowledge (Du Plooy-Cilliers, Davis & Bezuidenhout 2014). Scientific integrity must be maintained throughout the research and writing processes. Throughout the writing process, the researcher ensured that he acknowledged the use and recognition of work and ideas of others. Therefore, it was also important for the researcher to ensure that the work of other authors was not submitted word

for word as if it were the work of the researcher. That is, for any work cited and/or duplicated, referencing was done. This included acknowledging everyone who helped with this study.

Creswell and Creswell (2018) also note that it is unethical for a researcher to "cut and paste," or offer a written work that includes substantial passages from a single source without any changes. Some unethical practices related to literature used that should always be avoided include (Cresswell & Creswell 2018) avoid altering key words or phrases without proper citation, paraphrasing from different sources, self-plagiarism, combining cited sources with copied passages, and presenting a mix of copied material from multiple sources without proper citation, as well as citing non-existent or inaccurate information about sources.

In this study the researcher adhered to standards of scientific integrity, which include not inventing, altering, misrepresenting, or falsifying data (Dooley, Moore & Vallejo 2017). Additionally, the researcher refrained from swaying participant replies or manipulating the data to confirm her assumptions. This indicates that the research findings fairly represent the events that took place during the investigation. Scientific integrity, according to Akaranga and Makau (2016: 3), embodies a “commitment to intellectual honesty and personal responsibility for one's actions, as well as a range of practises that characterise responsible research conduct.” In this study, scientific integrity issues such as data fabrication and falsification, as well as plagiarism, were addressed.

3.9.4.1 Fabrication and falsification

Fabrication entails creating, inventing, and/or faking data or results, whereas falsification or fraud entails changing results or omitting some data or findings (Kour 2014). The researcher ensured honesty in reporting the collected data and the results of the investigation to maintain academic integrity.

3.9.4.2 Plagiarism

Plagiarism remains one of the most important aspects of higher learning institutions. It refers to the practice in which researchers are supposed to ensure originality and acknowledge where the information used was obtained from academic sources (Akaranga & Makau 2016). This meant that the researcher refrained from plagiarising other works by acknowledging the sources consulted by proper citation and referencing. The study was subjected to Turnitin software to verify whether it was compliant with the similarity index as specified by the

university. The researcher then submitted the Turnitin report together with the written submissions.

3.10 EVALUATION OF RESEARCH METHODOLOGY

Researchers are tasked with evaluating applied research methodologies to identify both their strengths and their weaknesses within the study (Creswell 2014). Leedy and Ormrod (2013) emphasise that research methods are inherently imperfect, casting continuous doubt on research findings. Creswell (2014) asserts that the purpose of evaluating research methodologies extends beyond informing fellow researchers about study challenges; it also contributes to enhancing the broader landscape of social research practices. Consequently, research methods must undergo an evaluation to elucidate the information sought, the collection methods, and the subsequent analysis. Kenneavy and Harnos (2023) further reinforce the necessity of evaluating data collection and analysis procedures to identify study limitations.

This study used a qualitative research approach that included interviews, document analysis, and observation for data collection. Creswell and Creswell (2018) emphasise that the primary strength of the qualitative approach lies in its capacity to delve deeply into explorations and produce detailed descriptions. This study was based primarily on interviews conducted with knowledgeable participants in the outreach, client services, special projects, sound preservation unit, and film preservation unit. This inclusive approach was essential since oral history content exists in diverse formats, including audio and video.

However, it is worth noting that although there were participants from the outreach, client services, and special projects unit of NARSSA, only participants related to information system management within NARSSA were interviewed. This selective approach stemmed from referrals by NFVSA participants and the study's focus on metadata and access to oral history. Given that NARSSA managed the metadata creation and access information system as the parent organisation, these specific interviews were deemed the most relevant. Some interviews had to be conducted online due to difficulties in arranging in-person meetings. These challenges arose from frequent work-related postponements, but the impact on data collection was mitigated by pre-interview discussions that yielded valuable insights. In addition, participants facilitated contact with their assistants directly involved in oral history content and the NAROS, the oral history content information system.

During data collection, a key participant at management level, to whom the researcher had been referred, had left the NFVSA, complicating efforts to secure an interview with the newly appointed participant. Consequently, the reliance on document analysis became crucial and some participants contributed insight into legislative aspects of metadata schema and oral history issues.

3.11 SUMMARY

The chapter provided a comprehensive overview of the research methodology employed in this study. It delved into various crucial aspects, including the research paradigm, research approach, research design, population under study, sampling methods, data collection, scientific integrity, and ethical considerations.

The chapter began with explanation of research paradigms with the rationale for the choice on interpretivism constructivism for this study, followed by research designs and qualitative approach and justification of the case study design. In line with the study objectives, a qualitative research method was adopted, population and sample were described as well as employing both document analysis and interviews as data collection techniques.

Furthermore, the section elaborated on the meticulous measures taken to guarantee that the collected data were subjected to ethical analysis and reporting to ensure that ethical standards were upheld during data analysis and reporting. Lastly, the chapter outlined the systematic strategies implemented to uphold the overall quality and rigour of the study. Measures to establish validity, reliability, objectivity, credibility, transferability, and conformability were meticulously addressed to enhance the study's overall robustness and value. These strategies boost confidence in the accuracy and applicability of the findings. The last section in this chapter evaluated the research methodology followed in this study.

The next chapter focuses on the analysis and presentation of the collected data, offering insights into the outcomes and implications derived from the meticulous methodology outlined in this chapter.

CHAPTER FOUR
DATA ANALYSIS AND PRESENTATION

4.1 INTRODUCTION

The previous chapter expounded on issues pertaining to the methodology employed. It focused on the research paradigm, research approach, research design, population and sampling, data collection methods, and instruments used, ethical considerations, and explaining measures to ensure the quality and accountability of the study. This chapter focuses on the analysis and presentation of data collected from document analysis and interviews with NFVSA participants. Primary data analysis is the original analysis of data collected for a research study. Analysing primary data is the process of making sense of the collected data to answer research questions, or to support or reject research hypotheses that a study is originally designed to assess (Kim & Liu 2017). According to Dudovskiy (2018), research data analysis is a process used by researchers to reduce data into a narrative and interpreting it to extract insights. The data analysis process helps to reduce a large amount of data into more manageable, coherent chunks. This study used qualitative methods to investigate the metadata schema of oral history content in South Africa to propose a framework to ensure access and retrievability. The following section describes the data analysis procedures used in this study.

4.2 DATA ANALYSIS AND PRESENTATION

Data analysis ensures that the reader understands theories and concepts used in research (Babbie 2013). To achieve the aim and objectives of the research, data analysis in qualitative research involves finding common patterns in the responses and critically analysing them (Dudovskiy 2018). Creswell and Poth (2016) state that conceptualisation occurs in two layers: the first basic layer is the more general data analysis procedure, and the second, more advanced analysis procedure embedded into qualitative designs. This suggests that organising, describing, elucidating, understanding, and accounting for data is the important part of data analysis. Additionally, interpreting data in terms of the participants' definitions of the situation, among whom the researcher becomes one to identify trends, themes, categories, and regularities, all of which the qualitative researcher is responsible for (Creswell & Creswell 2018). The current researcher attempted to interpret, make sense of, and give meaning to, the metadata schema for oral history content and build a framework to support this by organising, reducing, and describing the data.

This section begins with an overview of the study's participants.

4.2.1 Study participants

Qualitative researchers should focus less on the representativeness of the sample and more on how the sample or small collection of cases, units, or activities illuminates social life is chosen (Moser & Korstjens 2018). As stated above, the guiding principle in qualitative research is to sample until data saturation is reached. After collecting data from eight key participants in community projects and an information management system, where oral history content was collected and metadata were created, the current study reached saturation. However, despite having 16 participants, any participant who did not respond to or repeated an answer was omitted from the data presentation. Johnson and Christensen (2014) recommend that at least 5 to 25 participants should be interviewed, with the length of the interviews dependent on when saturation is reached.

The NFVSA is divided into six sections: legal deposit; sound preservation; film preservation; admin and registry; outreach, client services and special projects; and related materials. The outreach unit is in charge of oral history and the sound preservation unit is involved in some projects. Table 4.1 shows the breakdown of personnel at the time the data were collected.

Table 4.0.1: Breakdown of staff at the NFVSA

Unit	**No. of Staff**
Head of National Film, Video, and Sound Archives = 1	
Legal deposit unit	*2 permanent staff members*
Sound preservation unit	*1 permanent staff member + 1 intern (temporary placement)*
Film preservation unit	*2 permanent staff members + 2 interns (temporary placement)*
Admin and registry	*1 contract official (temporary placement)*
Outreach, client services, and special projects	*3 permanent staff members*
Information system management & arrangement and description unit	*3*

Total: 16 staff members

It is worth noting that most participants had some responsibilities beyond each section due to the limited number of staff, and this reflected a broad range of responsibility. An oral history project coordinator, for example, was also involved in the metadata of those records. Given that the NFVSA is a government institution established by the NARSSA Act and governed by NARSSA, it was not surprising that the information system section was managed by NARSSA, and some participants from NARSSA were interviewed for this study. Given the autonomy and structure of the NFVSA, it was not necessary to specify and link each participant to the role they played, as doing so would jeopardise the anonymity and confidentiality of some participants who chose to remain anonymous. For instance, by indicating that Participant A was a sound preservationist, the participant would still be identifiable because the section only had one staff member. As a result, for the purposes of data presentation, participants were referred to using letters of the alphabet, such as Participant A without role specification.

4.2.2 Data analysis

Data analysis is "central to credible qualitative research" (Maguire & Delahunt 2017:3351). Regardless of the method used to collect the data, data analysis begins with data editing, and the information gathered is referred to as raw data. According to Creswell and Creswell (2018), the first step in data processing is editing, which involves ensuring data is clean and free from inconsistencies and incompleteness. This process minimises errors, incompleteness, misclassification, and gaps in information obtained from respondents (Creswell & Creswell 2018). However, Gläser and Laudel (2013) state that there is no one single or correct technique to analyse and display qualitative data; how one does it should abide by appropriateness for purpose.

This study adopted a thematic data analysis. Thematic analysis can use interview transcripts, newspaper articles, questionnaire answers, diaries, films, photos, and field observations. It is a data analysis strategy, which is a widely employed method in all qualitative designs. Thematic analysis commonly used in identifying, analysing, and reporting patterns (themes) within data. The thematic data analysis is not simple and includes (Wellington 2015:263):

> "(a) dividing the data into 'units of meaning'; (b) classifying and grouping the units of meaning; (c) including new units of data into these groupings/categories; (d) searching for categories that are similar and/or that can be merged into a single category; (e) reviewing categories that contain large amounts of data to see if they can be split into smaller categories; (f) checking that the categories include all the data and are mutually exclusive (though some data may appear in more than one category); and (g) looking for links, contrasts, and comparisons between categories (constant comparison)" (Creswell 2014).

Simply put, thematic analysis is concerned with "transforming raw data by searching, evaluating, recognising, coding, mapping, exploring, and describing patterns, trends, themes, and categories in the raw data, in order to interpret and provide their underlying meanings" (Ngulube 2015:133). According to the research purpose, the goal of this data analysis was to examine the various elements of the captured data to clarify concepts and interpretation of the qualitative data into manageable categories, patterns, themes, and relationships (Bryman et al. 2014; Ngulube 2015).

To summarise, in the current study, qualitative data included recorded responses from document analysis and semi-structured interviews of NFVSA participants. Following data collection through interviews and document analysis, the data were analysed and presented to draw conclusions about the research objectives of the study. Data were collected in the form of text from document analysis, interview transcripts, and audio recordings obtained from interviews, all of which were analysed.

Qualitative data collected from the metadata records creation policy and/oy guidelines are analysed and integrated throughout data presentation. Notes from these documents were documented manually. The purpose of the study was to analyse metadata schema used for oral history content. Data such as the metadata elements and fields, as well as guidelines on access to oral history was collected. In the case of interviews, the researcher reviewed the interview schedule, listened to the audio recording, and read the transcriptions for any gaps or inconsistencies. In some cases, the researcher was able to recall the responses of the participants and confirm them with audio recordings (Creswell & Creswell 2018). The following section focuses on the presentation of the data in relation to the purpose of the study.

4.3 DATA PRESENTATION

It is important to remember that, in order to minimise bias, researchers are supposed to think more like observers than presenters when presenting data (Leedy & Ormrod 2013; Saldana 2013). In this study, the researcher was guided by a conceptual framework, but also "shaped the analysis and presentation of the findings to some extent by the researcher's point of view, disciplinary knowledge, and epistemology" (Braun & Clarke 2013:175). In view of the qualitative nature of this study, the data were condensed without losing meaning. This was supported by the fact that qualitative research is commonly used to generate deep data (inductive) rather than statistical data (Creswell 2014). The presentation of the data is organised according to the objectives of this study for the purpose of logically arranging the collected data. Data are presented verbatim, as well as through screenshots. The objectives of this study were to:

- describe the metadata schemas used for oral history at the NFVSA
- analyse standards for classification of oral history content at the NFVSA
- determine access to oral history content in the NFVSA
- determine technologies used for the metadata classification schema of oral history content in the NFVSA
- propose a metadata schema framework for oral history content.

4.3.1 Metadata schema used for oral history content

The first objective of this study was to describe the metadata schema used for oral history at the NFVSA. To ensure discoverability and access to oral history content, a metadata schema is essential (Fraser & Halvey 2017). The conceptual framework of this study was guided by metadata categories and metadata creation that perform accessibility and discoverability, as well as metadata elements applicable to the description of oral history content. The oral history content recordings are not easily searchable, making them unique among archival resources. This is because when creating metadata for oral history content, significant gaps may be discovered between what a metadata schema in use provides and what the oral history collection requires. The researcher intended to learn about the different types of metadata categories and metadata elements used for oral history content in this objective. On the other hand, this objective sought to establish metadata elements that may or may not be relevant and/or applicable in the creation of metadata for oral history content. That would also have

helped to determine whether there was a need to add other metadata elements not included in the current metadata schema and/or review the current schema. This section is divided into various elements of the objectives for logical data analysis. To align data with the objectives of this study, it was necessary to begin data presentation by determining the familiarity and knowledge of the metadata schema used by the participants.

4.3.1.1 Knowledge for metadata schema

It was important for this study to establish the familiarity of participants with the metadata schema used in the NFVSA because the analysis of websites (NARSSA) revealed that it was newly introduced and in the implementation stage. This information was critical because knowledge of the schema would allow participants to respond to the questions asked and, more importantly, share their experiences, since this was a qualitative study that relied more on participants' real-world experiences and the meanings they ascribed to such situations (Silverman 2020).

The researcher began by identifying the qualifications of the participant. The results revealed that all of the 15 participants held qualifications from post-matric certificates up to master's level at the time of data collection. For example, one of the participants indicated that he had a degree as well as short courses on archives, while the other indicated that he held a Bachelor of Information Science degree. On the other hand, one participant had obtained a master's degree shortly before the research. In addition to educational qualifications, it was equally important to assess the number of years of experience at the organisation. The experience of participants ranged from one year and six months to more than 25 years of service. This also assisted in determining the familiarity as it relates to the period on interacting with either oral history content or metadata schema and/or both.

When asked about how familiar they were with the metadata schema, participants reported being familiar with the schema in use, as they have used it, while others had never used it but have been exposed to it. This was because despite the metadata schema being officially rolled out and implemented at NARSSA in general, this was not fully the case at the NFVSA in due to network challenges. Participant A stated that, "*We were introduced and even trained to use the new metadata schema, but we had not created any metadata records because the system is always on and off due to network problems*." Participant B added that, "*We are very much familiar with the schema as I had created the metadata records even though it is slow due to*

system downtimes." All participants were familiar with the metadata schema in use because they had either interacted with the schema or had been introduced to the it as part of NFVSA activities.

4.3.1.2 Metadata schema used at the NFVSA

Participants were asked to specify the metadata schema used for metadata creation at the NFVSA. This question was a follow-up to identify what the metadata schema in use was used since it was indicated that the NARSSA employed a new metadata schema, but it was not rolled out to the NFVSA due to a network issue. All participants agreed that AtoM was used for metadata creation of oral history as part on the NARSSA collection. This was used because it was adopted as a national schema for South African public archival institutions. Participants acknowledged that they sometimes still used the Dublin Core because it could be used manually. Participant A acknowledged that the metadata in use (AtoM) was officially adopted and used by the archival institution under study. Participant E indicated that "*The metadata schema in use was AtoM*," but further stated that, "*I never used the new metadata schema (AtoM) but the archive(the NFVSA*) *is in the process of migrating metadata from the Dublin Core into AtoM.*" This showed that the metadata schema in use during the time of data collection was AtoM.

Since participants were familiar with the metadata schema, it was also necessary to identify metadata elements and fields used for metadata creation for oral history content. It was thus relevant to ask the participants whether they had identified which of the metadata elements were applicable in metadata creation of oral history content.

4.3.1.3 Metadata elements for oral history content

The conceptual framework in this study was used to identify elements applicable to metadata creation of oral history content, as well as elements that were not covered by the metadata schema but would have been included in oral history content metadata. As already indicated, the metadata schema used had just been implemented, as such the participants had experience with the two schemas, namely AtoM and Dublin Core. The researcher also hoped to learn about the participants' experiences with AtoM, as the majority of them had used Dublin Core prior to the implementation of AtoM. The researcher had hoped that previous experience of the participants would have improved their understanding of the relevant metadata for oral history

content. Participants were asked to explain the elements that were relevant and used for the creation of metadata for oral history content at the NFVSA. Furthermore, participant G stated that they could not comment on the metadata schema (AtoM) because they had hardly created any metadata since its implementation and indicated that they needed more experience with the schema. While participants C, D, F, and H indicated that all elements were applicable to oral history content metadata, participants A, B, and E provided contradictory responses. These participants agreed that when using the Dublin Core schema to capture metadata for archival materials, including oral history content, all elements were used, which was not the case with AtoM. Participant A indicated that there were some descriptions that did not fit in the said metadata schema. Participant B stated that "*It was almost impossible to record the physical description of an audio record*." Participants E added that even though they created some metadata records, they did not use all the elements to create metadata for oral history content. This was confirmed by the document analysis, as the researcher also discovered some metadata elements that were not filled with information, despite the metadata elements being applicable to oral history content records.

When asked why they hardly used or did not use some elements in their metadata creation, they stated that they were still learning the system and taking a long time to create records due to internet connectivity issues. For example, Participant B stated that "*Dublin Core did not cover some of the metadata information that was supposed to be covered, but I cannot comment on AtoM. AtoM has more elements than Dublin Core, but I am unsure of the elements that might be used or not, as I never created metadata from AtoM.* This was because we always experience network connectivity challenges, and more often there will be system downtime."

Furthermore, Participant E indicated that: "*Since we migrated to the new metadata schema, I have not used the schema very frequently, and as such I am not satisfied with what it is. I still need more experience with it to know which elements are applicable or not. What I know is that all records that were created using Dublin Core were migrated to AtoM. And that some of the Dublin Core metadata were still applicable with the current AtoM."*

Participants acknowledged that they had a minimally encountered the metadata system but they could still identify the elements used and/or not used to create metadata for oral history. Internet connectivity arose as the major reason for either not having created metadata and/or having used the metadata less frequently. Both those with experience and those without experience

with the metadata schema concurred that they needed training in the newly implemented metadata.

Participant F acknowledged that: "*There are some system downtime issues, but the main reason for the lack of use of the new metadata schema is that they are still training archivists on the metadata schema. Some started using the new metadata schema a month ago after full-fledged training."* (This was a month before the data collection period).

Participants were also asked to confirm whether there were metadata elements that could not be used to create metadata for oral records. Participant F responded by stating: "*There has been a lot of data cleaning during the migration from the old NAAIRS to new NAAIRS. We had added a lot of information to the records that was not there in the old NAAIRS records."*

Participant F further stated that: "*Metadata schema was changed because it was no longer compatible with the creation of metadata by the archival institutions. Even NARSSA had to change the metadata system to address the gaps and missing elements of the previous metadata schema. There was data from oral history content that would not be described using the current metadata schema."*

From the analysis of the NAAIRS database, the researcher identified the following element and/or fields:

- archival repository/depot: archival documents are kept as entities or sources in accordance with the office that developed or received them.
- physical location identifiers (reference numbers, file, volume, and box number).
- Classification system
- part number (the number of the part in cases where multiple files have the same reference number);
- description (a brief synopsis of the archive unit is provided, primarily containing the file's original office title);
- starting and ending time (the item's final dates over a period of time);
- Remarks (detailed information, including cross-references),
- Summary (only information related to national non-public record registers). A synopsis of a collection's contents, such as an explanation of the kind or format of the collection for audiovisual and manuscript entries.

Both remarks and summary elements were intended for the archivists' notes.The results indicated that the metadata schema used was not compatible, as the participants referred to gaps and missing elements in the system. It is for this reason that participants were asked to explain the metadata creation for oral history content.

4.3.1.4 Metadata creation process for oral history content

It was critical to understand the metadata creation process to determine whether the metadata schema used was the best fit for oral history content metadata records. The study examined the metadata schema system and the metadata creation process. This was also significant given the unique nature of the oral history content. The document analysis results indicated the following:

- Metadata standards were embedded in the metadata schema.
- There was a guideline on the scope of each element, which included what information should be covered within the specific element.
- Data templates that related to describing metadata standards but could not be edited and/or customised were provided
- These templates, based on descriptive standards, could add or remove fields without custom development, which was also not possible.
- The schema allowed either single and/or multilinking descriptions to authority records and accession records in the data entry templates.

Participants were also asked to share their experience with metadata creation to enhance data collected from document analysis. This was also linked to questions about how familiar participants were with the metadata schema, as well as how applicable and/or not applicable the elements they would have used during the metadata creation process for oral history content were. It was critical for this study because the conceptual framework aligned with the research purpose of this study, which was to provide a detailed description of metadata schema for oral history content. Participants confirmed data from document analysis. It was revealed that when creating metadata, they could not adapt the metadata elements according to the needs of the content described. This question was also open, so participants could respond from their own point of view, focusing on various issues of experience. In general, participants agreed that the metadata schema in use contained more elements than the previously used one. Furthermore, "*the metadata schema in question was less user-friendly, while creating a record for oral*

history in itself is difficult," lamented Participant D. Participant A stated that: "*as it took us such a long time to ensure that all of the necessary information were captured.*" The necessary information to be captured referred to all data required for metadata creation of oral history content. Participant C contradicted Participant D and explained that: *"Metadata creation has been simplified with the use of the metadata schema in use, as it allows us to capture all the necessary data on audio-visual records. The only challenge is system failure and connectivity, since the new schema requires the Internet connectivity to function."*

The researcher also examined the metadata records of oral history content. The researcher expressed concern that some other elements had not been completed. Participants were then asked to comment on the quality of the metadata created for oral history content.

Participants emphasised that they strove for metadata quality and were guided by the elements available in the schema. Participant E acknowledged that some information was sometimes missing due to information unavailability from the time when the oral history content was collected. Participant B stated that: *"As much as sometimes do experience gaps in our metadata records, it does not make the record the poor quality one. The information provided for is always correct and relevant metadata elements are used. It does happen but on a rare occasion the oral history team missed some data relevant for metadata creation while on the field."*

Participant G also indicated that: *"We always strive for accuracy, provenance, and consistency in metadata creation process. We also ensure that our metadata conforms with the metadata standards in use. Metadata created are to the most part a true representation of the oral history content."*

Participants confirmed metadata quality in metadata creation despite the fact that the templates available and aligned with specific standards did not provide clear instructions on how to enter certain fields. As a result, it was necessary to follow up on this question by enquiring about the metadata categories covered in the creation of metadata for oral history content.

4.3.1.5 Metadata categories recorded

To ensure access to oral history content, various metadata categories can be created. As a result, participants were asked why they created metadata in the first place. The objective was to identify the type and/or categories of metadata created at the NFVSA for oral history. Administrative, descriptive, and structural metadata were among the types of metadata created

at the NFVSA. Participant F stated that: "We *create metadata for administrative and descriptive purposes, as well as to show components of records. There is also an area for the administrative outline where a bit of background is provided on the particulars and that the elements are self-explanatory and as such straightforward."* This did not explain why some elements were not completed despite being straightforward. One would have anticipated that the relevant metadata elements were completed to ensure metadata access to oral history content. However, all participants confirmed creating administrative, descriptive, and structural metadata. Participant A added that "Yes, *as we do metadata, we also need to do preservation metadata."* The researcher could not ask the participant to further explain the process for preservation metadata simply because the participant stated that they have not yet created preservation, but most importantly, preservation metadata were not the scope of this study.

The researcher also confirmed that all metadata categories were covered in the metadata record through document analysis. The participants mentioned the metadata creation process and confirmed that all metadata categories were covered in the process. They also specified that all elements were applicable in both the previous metadata schema (Dublin Core metadata) and the current schema (AtoM).

In relation to the description of metadata schema for oral history content, the other objective of this was to analyse the metadata standards for oral history content in order to gain a better understanding of the metadata process and creation for access. As a result, the next section presents data collected for that purpose.

4.3.2 Metadata standards used for the classification of oral history content

The second objective of this study was to analyse the metadata standards for classification and access to oral history content. To ensure access and discoverability of archival materials, various metadata standards were used. Grimsley and Wynne (2009) emphasise that the importance of metadata standards for oral histories is to improve access to oral history materials. The conceptual framework for this study advocated for successful metadata for the accessibility and discoverability of the oral history content of the NFVSA. Metadata standards are essential for categorising and classifying oral history content. However, Ribeiro (2014) explains that using classification to physically arrange documents and give them a physical and systematic organisation is not done at all in historical archives, because the main concern is the

utilisation of space in the stacks and because the fonds are not freely accessible. Despite this, the main purpose of classification is to organise and categorise information and materials, resulting in greater accessibility. It was important for the researcher to understand the role of classification in archival institutions and the standards that might be relevant in, and applicable to, oral history content. The quality of classification depends on the knowledge of standards used. It was thus important to ask a question on the knowledge of metadata standards used for classification in general and oral history content in particular. The following section begins with knowledge of metadata standards by participants.

4.3.2.1 Knowledge of the metadata standards

Metadata standards are essential for creating metadata because they provide guidelines for the metadata process in organising and linking oral history content to each other, but also to other related information and/information sources in the databases. These metadata standards, on the other hand, aid in describing and categorising/classifying materials for easy access. Researcher noted that no specific metadata classification standards were used because there was no need to classify the content in their various formats, and that an in-house classification was mostly followed (Fell & Ou 2022). Fell and Ou (2022) further argue that there is no need for such classification, as archival materials are not put on shelves to be accessed openly anywhere. It was thus appropriate to ask about the participants' knowledge and use of metadata standards in their work.

It was not surprising to find out that participants referred to the ICA standards (ISAD(G), ISDIAH, and ISAAR) standards embedded in the metadata schema in their responses, when asked about standards to classify oral history content. According to Participant F: "*We use the ICA standards, which are helpful, as the instructions are clearly outlined, even though it is not always done properly by archivists, as some are still being trained properly as they adapt to AtoM. I think records managers and users are not properly trained to identify the various elements involved in the metadata creation of oral history content or archival materials, in general. But oral histories are audio-visual records, and we describe a lot of audio-visual content.*"

Participant G commented in general about the metadata used in archival institutions. From the interview, the researcher saw that the participants did not really have knowledge about the metadata standards because even when asked several times to clarify the statement, participants

kept on giving general responses. Participant H clearly indicated that he did not know much about metadata standards. Participants A and B agreed that they were aware of the standards, but did not always refer to them because they were simple to apply and remember. This could be linked to participants' knowledge of metadata schema, as these standards were embedded on the schema. Some participants indicated that they were familiar with the metadata schema while some were not.

Participant A mentioned that ISO metadata supports the metadata standards for classification of oral history content, but could not specify the ones used for classification purposes. "*As archivists, we have ISO standards to comply with, as they guide us on what is expected of us in creating the metadata. But the ICA standards for archiving are also very important. Especially the ISAD(G) that we use to capture archival descriptions."*

From document analysis, the researcher discovered that ISAD(G) was used more frequently than other standards available on the schema. Thus, it could be that ISAD(G) contained metadata elements applicable to archival institutions. However, the standards could not be used for classification of archival collections, including oral history content. As a result, in addition to the responses of the participants, a question was asked about metadata standards for the classification of oral history content. All participants clearly indicated that they had engaged in the standards in their work. The next section's question was about classifying oral history content.

4.3.2.2 Classification standards of oral history content

Metadata standards can help users find and access the resources they need by providing standardised methods of describing and categorising resources (Raghavan & Krishnamurthy 2013). The classification and/or categorising of archival materials is an important tool for ensuring that users can discover and access oral history content. To understand the importance of access to oral history content, it would be beneficial to have a better understanding of basic classification standards. Participants were asked to explain the classification process and the standards used to classify oral history content. Knowledge of classification and its processes is required to analyse metadata standards. However, archival classification works differently from bibliographical classification in information centres such as libraries. In the case of archives, classification is also a means for organising information and knowledge, but it primarily enhances the context of the production and accumulation of records dictated by the

functions, activities, and organisational structure of the creator. As a result, participants were asked about archival material classification in general and oral history in particular. The results showed that there was no classification of oral history content and following were participants' responses.

Participant A stated that: *"In archives, we do not classify like libraries do. We do not allocate classification numbers but topics and/or title showing content. We classify the content according to topics and subtopics. Secondly, we follow the file plan classification scheme. This is because in archives we do not classify individual sources but in groups. For example, oral history collection will be in multiple formats and multiple topics. If we classify according to the notation system, these sources will be separated."*

Participant E stated that: *"In archives, the emphasis is on the principle of provenance and original order. We are unable to change the description provided by the original creator. If you do not keep the materials in their original description and order, they will no longer be useful or accessible. Furthermore, if we do not keep the original order, we will lose the links between identical or related items."*

The researcher probed further on the issue of file plan and classification according to topics. It was clear from the metadata system that one of the fields was named 'Content field'. This field was meant for classification and subject organisation of archival collections. The researcher also noticed a metadata field to describe the items' content. In the 'Content' metadata field, archivists and/or metadata practitioners had to specify the keywords and concepts that specify the subject content and aboutness of the sources, thus classification. In addition, there was a system (classification system) filed wherein archivists had to specify the classification system used and the reference. In this system field, that was where the file plan was indicated. The file plan is a pre-determined classification plan by which records are filed and/or electronically indexed to facilitate efficient retrieval and disposal of records (NARSSA 2016).

A follow-up question was asked about whether the file plan system and/or the ISAD(G) helped to determine where to classify the materials. This question was also posed because participants kept referring to the file plan and the 'Content field' as one of the metadata elements in the metadata schema that helped with classification and subject organisation. Participant E stated that: "When *creating metadata, we research to gather more information about the archival*

group so that we are aware where the content can be found, and that is mentioned in the 'Note element'." That referred to classification from the ISAD(G) field.

Participant E reiterated what Participant A indicated: *"The principle of provenance is important in archives, not in the organisation of the subject. Oral history is multidisciplinary and, as such, we cannot classify it anywhere. But we have a file plan to organise and describe archive materials."* Participant A, on the other hand, emphasised the use of the record management policy to describe and classify archival materials. In addition, Participant H highlighted the importance of the file plan as a classification system, citing the fact that the "*file plan has been approved as a classification system in archives."* The participant added that: *"there is a legislative framework that guides everything we do at NFVSA, including the creation of metadata and the system we are to use."*

The mention of the record management policy and legislative framework for classification purposes raised the importance of guidelines in metadata standards, thus classification of oral history content. It is for this reason that a question was asked whether ISAD(G) provided clear guidelines on how to determine the keywords to use in describing the content of these materials. The following section presents responses to the classification guidelines to determine and select keywords for description.

4.3.2.3 Classification guidelines for content specification of archive materials

The quality of metadata helps to manage and access oral history content in an effective and efficient manner. The quality of classification would also be determined by the specification of how to do it, as well as the guidelines for describing the content (Yang, Chen, Feng, Yan, Zheng & Zhang 2020). The purpose of this study was to examine the use of metadata standards to classify oral history content. The analysis is based on whether the standards provide clear guidelines to that effect. It was clear that the classification standards used were ISAD(G) and/or file plan, which were stated to align with organisational functions. Depending on the origin of oral history content, it is sometimes difficult to classify oral history content, especially when it did not belong to any organisation/department. The participants would rely on the "Content" field from the metadata schema and standards. Participants were asked whether there were guidelines on how classification and keywords should be created and they had to explain the guidelines in line with the metadata standards, as supported by the record management policy.

Participant A stated that ISAD(G) *"provides guidelines on what information to add and/or to include while creating metadata for archival materials, including oral history*." Participant C explained how metadata standards helped to *"identity keywords so that funds are linked to each other*." Participant D added that metadata schema "*provides an opportunity for archivists to make notes and indicate where else they can find information or a source related to the one in question."*

There is a metadata element that allows archivists to enter keywords, but the participants indicated that there was no guideline on assigning those keywords. For example, no vocabulary control was used, such as a thesaurus to ensure consistency in assigning keywords and vocabulary used.

Under the context area, Participant F indicated, *"First we create the authority record, which is the Office of Origin, and we then create the fonts or capture the fonts, it is important that we link it to the Office of Origin, and here the archivists are expected to conduct a little bit of research on the archival group that they are conducting so that they can have the biographical history. There is also an area for administrative outline where a bit of background is provided on the particular archival group."*

The researcher examined the records, focusing on the 'content element,' which was discovered to be more of a classification element. The results revealed that the captured keywords were not intended for the specification of the subject, but for giving more details about administrative metadata, as the title was always used and/or broken into keywords. The keywords did not really address the content of the collection at hand. In some cases, no information about archival group was written. Simply put, there were missing data and gaps on the 'Content' field, which was meant to include assigned keywords for classification of the oral history content.

Given the challenges associated with classification and/or content specification, participants were asked which metadata standards they would prefer to use for classifying oral history content in the NFVSA. The conceptual framework was designed to determine whether the classification standards were used for the oral history content. This raised the question of what factors influenced the selection of standards or schema in institutions. This was also due to the participants indicating that they switched from one standard to another.

4.3.2.4 Choice of metadata standards

It was important to determine what informed the choice of metadata standards used at the NFVSA. The intention was that participants would provide reasons for their choice, allowing for extensive analysis of the standards, which was the objective of this study. As a result, the researcher wanted to know if and how they were involved in the selection of the metadata standards they used for oral history content, as well as what factors that influenced their decision. This would have determined whether the participants were satisfied with the archival institutions' efforts to establish international standards and adhere to best practises relating to classification of oral history content. The participants' expressions suggested that they would like to have another standard, if possible. One of the participants expressed a sense of compliance rather than satisfactorily completing the work in order to achieve the desired outcome. The researcher broadened the question to enquire whether the participant would have considered principles of metadata standards used by the NFVSA, if they were given an opportunity to make a choice. The researcher had to elaborate on the principles, as some participants seemed not understand the principles guiding metadata choice. Participants agreed that they had no significant role in the decision about which metadata standards were in use at the NFVSA at the time. Participant F stated that they did not directly play a key role, but as part of monitoring and evaluating metadata and related activities, they identified issues with the then current metadata schema and standards. "*Our institutions follow certain policies and legislation for any other activities, including the purchase and subscription of resources.*"

This was a first step towards identifying key legislation governing metadata and metadata access to archival materials. According to the results of the researcher's document analysis, the legislation was not specific to oral history content but to archival materials in general. Participants agreed that metadata are important for accessing archival materials, including oral history content, despite their lack of specificity in relation to legislation and metadata. Despite that, metadata remains the most important tool for access to, and discoverability of, oral history content. Hence, the next section focuses on the objective on access to oral history content.

4.3.3 Access to oral history content at the NFVSA

The conceptual framework sought to determine access to oral history by describing how oral history content was made available at the NFVSA. Access to oral history content is important given the significant role it has played and continues to play in societies. Scholars have varying

views on the importance of oral history and oral traditions. For example, Ngoepe (2019) emphasises the value of oral sources in reconstructing and reinterpreting history, especially for people whose stories have never been given the opportunity to be told. Bhebhe and Ngoepe (2021) add that oral history is a vehicle for promoting social justice and for accessing the long-neglected history. People, events, and experiences that have shaped personal lives, communities, and national identity are included in various ways. Oral history content is made available through oral history. Schellnack-Kelly (2017) emphasises that oral history is significant to create an understanding of incorporating indigenous knowledge, the myths, legends, and experiences of all communities as evidence of the necessity to preserve wilderness areas for the benefit of all communities and ensuring their accessibility to all. Bhebhe (2015) indicates that oral history can be used to promote cultural development and increase community knowledge of social concerns that have an impact on people's daily lives. According to Bhebhe (2015), oral history has the potential to capture and preserve community people' viewpoints in their authentic and natural environments. settings. When oral history content is recorded, what should follow is accessibility. Access to oral history content is critical to identifying, recognising, and respecting various aspects of heritage (SAHO 2019).

Access to archival material is one of the main services offered by national archives. Although oral history content is intended to be used by researchers, many organizations—including the NFVSA—find it difficult to grant access. Many factors contribute to inadequate access, but the most important one is the absence of consensus or standardisation in classification methods (Wynne 2009). Metadata facilitates access to oral history content because the descriptions are searchable from any database. To establish accessibility of oral history content, the formats should be identified.

4.3.3.1 Formats of oral history content

Archives are of a mixed or hybrid nature, with records that are digitised and those that are born digital. However, some of these oral history contents were still available on audiocassette tapes, minidiscs, and mini-DVDs. This suggests that some oral history content is still in analogue form while others are born digital, according to Bilushchak and Fedushko (2016). This was the case at the NFVSA, where newer materials were collected in digital formats while the older ones were collected in older formats such as audiocassette tapes. Despite the formats (old or new), oral history content should be made accessible to users. The NFVSA used channels such

as finding aids (to be discussed later under 4.3.4.4 Finding aids and access) for accessibility and discoverability of oral history content that is available on their databases. Therefore, it was crucial for this study to describe the storage for oral history in relation to access at the NFVSA.

4.3.3.2 Storage for oral history content

The researcher asked the participants about storage for oral history content. Some participants were reluctant to answer the question and others indicated that they had not been to storage for a long time, as such they could not comment because they did not know. Participant A stated that "*I am always out on oral history projects, and that I am not responsible for preserving and storing the recordings. As such I am not sure what happens to their storage after I gave them to the person responsible for."* Participant A further indicated that: "*I am aware that at some stage there was some storage challenges but since we had some temporary and interns, a lot of progress had been made*.' Participant B and C also stated that they could not comment on the storage, as it was a different section, and as such they were not fully aware of what the status thereof was.

The researcher confirmed a variety of formats and storage of the oral history content during data collection on the site. The oral history content that has been captured since the early 2000s is hidden on tapes stored in boxes at the NFVSA (Ngoepe 2019; Mukwevho & Ngulube 2022). This suggested that those oral history collections were not yet described nor classified. This led to oral history that would not be accessible or discoverable by the public as people might not be aware of its availability and/or existence.

Participants agreed on the staff overload as the reason for the backlog at the organisation. The issue of staffing is elaborated on later in this study. This backlog should be addressed to ensure that oral history collections were fully described, thus allowing access thereof. The availability and/or non-availability of metadata records, and storage as such, has an impact on search strategies of the users at the NFVSA. The respondents were then asked about the users and their searches on oral history content.

4.3.3.3 Search strategies to access oral history content at the NFVSA

As the objective of this study was to analyse access to oral history content, it was also relevant to look at access from the users' side. Participants were asked about the experiences of their users while accessing oral history content. Participants were expected to reflect on the services

offered to users, and the queries that users pose in their endeavour to access those content. One of the questions asked was to explain the search strategies available and in use by users. Search strategies are important for the successful retrieval of information from any database. The right search strategy produces the right information access. Access to oral history content also requires effective search strategies and searches.

Participant F indicated that users were mostly concerned with not finding what they were searching for. From the NARSSA website, there was an explanation why users could sometimes not find the records. One reason for this was that the record one would be looking for has not been ingested into the system yet. In this case, users would have to utilise the manual inventories that available in the reading rooms of the various repositories. Another reason could be not using the relevant search terms. Most importantly, the retrieval of records in the NAAIRS was influenced by search terms. Depending on the scope of the information required, users may use single or multiple words. The system included search strategies for Boolean operators. From document analysis, the use of these operators was detailed on the NAAIRS database to guide the users on how they were used and what they meant.

Document analysis, on the other hand, revealed that users were free to formulate appropriate search words independently because metadata standards and the file plan for classification did not use any type of controlled and/or standardised list of words. Given this, participants were asked about the process of adding search terms to the NAAIRS. It was evident from the search that not any keywords and/or terms retrieved positive results. For example, Figure 4.1 below shows 117 results. Users might find it difficult to know which of the 117 addressed their information needs, and even where on the system to find more information about the results.

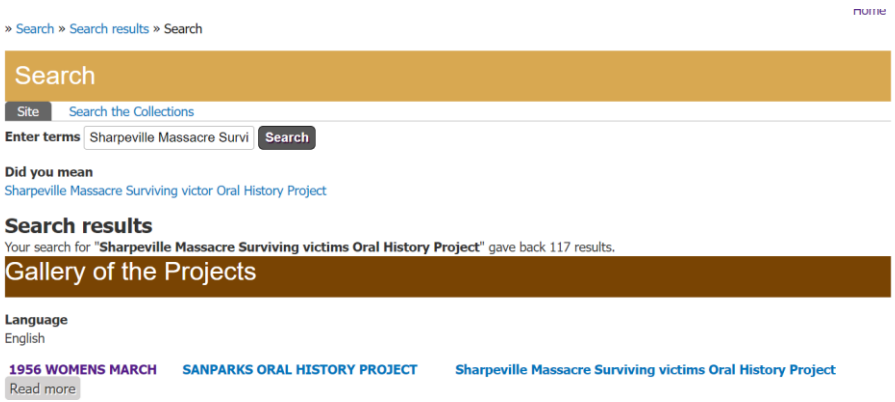


Figure 4.1: Search results for Sharpeville Massacre survivors

Participants C, D, and E emphasised the importance of provenance as the reason why they could not create subject entries for oral history content and/or archival materials, in general. "Provenance is an ordering principle comprising two dimensions: respect des fonds, the aim of which was to ensure that records originating from the same creator or creating body remained together, and original order, which aimed to ensure that records were kept in the order imposed by their creator" (Lapp 2023:123). In addition to the lack of search terms, the database search was not user-friendly.

However, the researcher shared the same sentiments with participant B, as during analysis of the database, it was found to have had limited search terms providing users with minimal search options. The search sometimes yielded 'no match found' when one participants used own term while, sometimes, they might retrieve incomplete information and/or results might show the results available but could not access the content. The researcher analysed the search process and Boolean operator search strategy. Even with the search strategy being used, no results were found. This related well with the classification process as explained under the classification guidelines section (4.3.2.3). This section indicated that there was no vocabulary control to ensure the keywords were used to search and retrieve archival groups. In some other instances, the results would indicate 'no access' to the records. Participants indicated that 'no access' would in most cases relate to legal and ethical issues for oral history content. As reported earlier, some oral history content is not usually put on open shelves for access, as there are ethical and legal issues relating to access to them. This prompted the researcher to investigate the issue of access to oral history and the ethical implications to that regard deeper.

4.3.3.4 Ethics and legal access to oral history content

When designing an oral history search system, consent, ethics, and control are of primary concern to a variety of stakeholders (Boyd & Larson 2014). As the objective of this study was to determine oral history access at the NFVSA, it was necessary to collect data on ethical issues related to oral history access. To determine access to oral history, the researcher examined documentation on borrowing oral history content, as well as data from interviews.

All participants agreed that some oral history content like any other archival records were only available for access with permission. As a result, participants were asked: "What *ethical obligations were to be considered when accessing oral history content?"* Participants mentioned that users were required to complete documentation for permission before

borrowing or accessing specific oral history content. The data collected from the document analysis revealed ethical and legal constraints regarding access to oral history content.

The participants also discussed permissions and ethical use of oral history content. According to Participant F, "*Even though oral history content is searchable and accessible through online systems, there are also ethical issues with access*."

Access to oral history content, on the other hand, was hampered by the absence of legislative framework detailing the exact permission, ethical procedures, and requirements available to the users. Participants agreed that the existing record management policy and the NARSSA Act did not hold direct accountability for oral history access. Participant F added: "*The policy available contains a general guideline for NARSSA and related archival institutions. It should be revised to include oral history content*." The researcher was able to analyse the documents and agreed with the participants that the NFVSA should reconsider the legislation governing access to oral history. In addition, there has been always a concern about ethical issues relating to access to information through technologies. Access to, and creation of, metadata for oral history content could be explained in terms of technology, as it relates to database and information systems, as well as access to oral history. As a result, the following session focuses on data collected on technologies for accessing and creating metadata for oral history content.

4.3.4 Technologies used for the metadata and access to oral history content

Archival institutions use many types of technologies. Understanding the technologies used for metadata around providing access to oral history content was also important, based on the participants' experiences. Literature demonstrates how technology has had a significant impact on how archival institutions provide their services and functions. The conceptual framework served as a foundation for identifying the technological tools used at the NFVSA for the metadata schema for oral history content. It was clear that technologies were used for the description of metadata content and the creation of metadata for archival materials, including oral history content.

To begin with, the metadata schema used at the NFVSA was web-based application that is open-source and based on standards describing and accessing archival materials. Second, there was an online search and access system for oral history content. Finally, the fact that NARSSA was in the process of "migrating more than 8.3 million entries" from the old NAAIRS database

to the new database (NARSSA 2016) when the data was being collected, indicated that the NFVSA had some metadata creation technologies. The study objective of this study was to determine technologies used for oral history content. This section is subdivided into the following subheadings: technologies for access to oral history content, technologies for metadata creation, and benefits of technologies used for oral history content.

4.3.4.1 Technologies for access to oral history content

The conceptual framework of this study sought to determine the technologies used for metadata creation and access to oral history content. To achieve this, it was important to begin with identifying technological tools available for accessing archival materials at the NFVSA.

In today's era, archival institutions are faced by the changing face of technologies. The use of technologies brought about change in access to oral history content. Database was already mentioned as a finding aid for access to oral history content. Participant F stated that: "*Oral history contents are fragile and complex to access as they are not easily browsable like other archival materials. Without the correct hardware and software, they will be not usable and accessible. These issues affect most the audio-visual formats of oral history content. We have some audiovisual formats which we can no longer access as they are not compatible with the new technologies in our organisation. Some collections are unprocessed collections while some are technically lost in the system, they are somewhere in the system but cannot be located."*

Technologies are changing rapidly, making the existing technologies and oral history content incompatible. This has a negative impact on access, as some of the collections are either hidden in unprocessed collections or may not be available for years, as the more the backlog, the more technology changes. This backlog affects metadata from the finding aids and availability of inventories. The next section provides a detailed description of finding aids at the NFVSA.

4.3.4.2 Finding aids and access to oral history content

The objective of the study was to determine access to oral history content. Finding aids were used to search for information in archival materials. The NAAIRS serves as a finding aid to assist users of archives to identify and locate archival materials that are relevant to their need. The NAAIRS contains only metadata pertaining to archival material rather than the document texts themselves. The NFVSA made the oral history database available for search and access.

The NAAIRS was the database from which archival materials were accessed. Prior to the interview with the participants, as part of document analysis, the researcher also searched the NAAIRS and in most cases found 'no inventory' or 'no access.' Data from the document analysis also showed that there were old and new NAAIRS that were both active as finding aids for access oral history content. The NARSSA landing page also indicated that one could search both the old and the new NAAIRS even though the new NAAIRS was still under construction (see Figure 4.2 showing the NARSSA website's landing page).

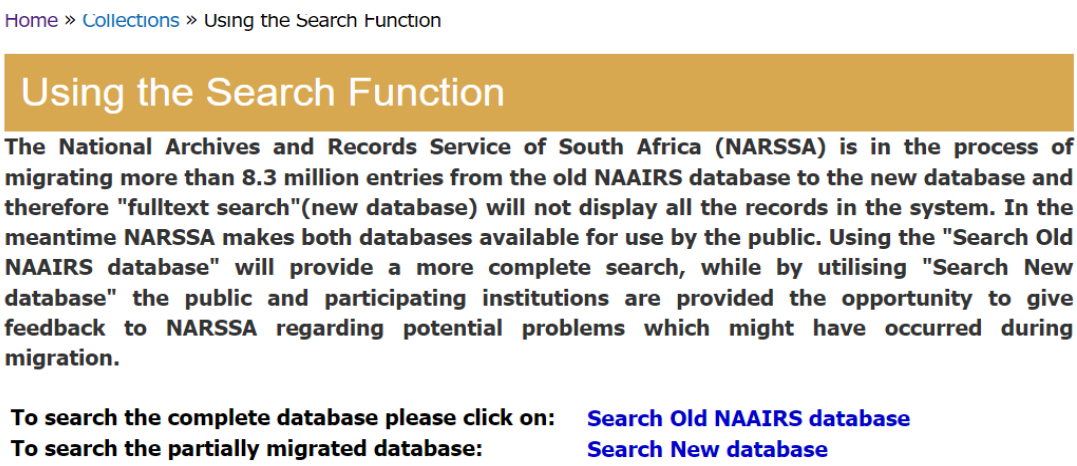


Figure 4.2: NAAIRS Search function: Old and/or new database

Despite the availability of the search function, some oral history content does not have inventories. Figure 4.3, for example, shows a search for public programming on oral history with the search results "No Inventories". Some participants attested to the absence of inventories in the finding aids. Participant D agreed that the system's content 'got lost' at times. *"This means that the files are somewhere on the system, but not discoverable or accessible. In some cases, due to the backlog, it may suggest that there are no metadata records for that specific content."*



Figure 4.3: Inventories for oral history content

As reflected in Figure 4.3, there are no search aids and inventories for oral history content online. Participant A stated that: "NAAIRS *is not searchable by the public, but users can always access the new website."* Participant B indicated that: *"the search function on the website is not easy to use."* Participant D stated that: "*I don't know exactly what happens when you search as a user, but I know that users always visit our reading room and we do searches for them. Most of them complained about how difficult our system is."*

According to Participant B, "*I acknowledge that the search function on the website is not easy to use for users. For now, what we do is search and upload to our website. The other confusion might be that there are two databases: the old NAAIRS and the new NAAIRS."*

Participants did not hold the same view with regard to the NAAIRS being searchable or not. For example, Participant E disagreed with participant B, and indicated that: *"NAAIRS is searchable for various terms, and there is also 'help search' information on the site. The site allows users to filter their search according to their needs. We comply with provenance but also add more information to the metadata fields to assist users."*

Participant A indicated that in a situation where archival materials are not accessible from metadata schema, archivists would download the collection an upload on the NARSSA website. This suggests that the NFVSA made an effort to use available technologies to make oral history content accessible. The results showed that the NAAIRS's lack of metadata and/or incomplete records contributed to oral history through the NAAIRS not being accessible.

Hence, the next section focuses on technologies used for metadata creation for oral history content.

4.3.4.3 Technologies for metadata creation at NFVSA

Metadata have significance for the access to oral history content. Additionally, technological tools are also crucial to ensure updated and accessible metadata records for easy access of oral history content. It was important to check technologies used for metadata creation for the oral history content. Given the technological nature of the metadata schema, metadata creation done was online.

Metadata schema used at the NFVSA is an open source system designed to transform access to oral history recordings by presenting digital audio or video in conjunction with a transcript and/or index and by connecting users to a particular moment in an interview via browsing or keyword searching. Participants agreed that metadata creation had been improved since the dawn of online metadata. They also acknowledged the available information management system and that even though there was still a backlog, technologies simplified some manual way of metadata creation.

Participant F stated that: "*Technologies have simplified arrangement and description of oral history content. Collections that were in paper form are now stored as electronic files. Managing metadata in an online space is far much better that the manual space.*"

Participant E further indicated that: "*There are digitisation tools. The fact that we still have archival materials in old format and some no longer compatible with new technologies, it means that the collections should be digitise for improved access and metadata creation. This is because even digitised and/or migrated document still require metadata, and thus can be easily manageable with the use of technologies."*

From the analysis of the databases, a significant percentage of the records in the NAAIRS are in Afrikaans, as are the archival materials to which they refer; therefore, retrieval on the basis of subjects may occasionally not be fully accomplished without the inclusion of the Afrikaans search terms. For example, 'Boek (book), tesis (thesis), pamflet (pamphlet) are found as search terms. These terms refer to archival materials in the custody of national and provincial archival repositories (NARSSA website). This could be the reason why some participants lamented about the need for translation as part of metadata creation. Participant C stated that: "*oral*

history collections are in most cases multilingual, and the system cannot accept some of the terms used. At the same time, it sometimes difficult to analyse the documents as some are in foreign language such as Dutch." Participant D added that: *"The issue of language plays a role in metadata backlog"*.

In relation to technologies, another issue was network connectivity, which also affected metadata creation. During network challenges, no metadata would be created, as the metadata schema in use was a web-based system. Despite the challenges relating to metadata creation using technologies, it also has benefits, which can be related to the use of technologies for oral history content. Hence, the next section focuses on the benefits of technologies used at the NFVSA.

4.3.4.4 Benefits of the technologies used at NFVSA

It was critical for this study to establish the perspectives of the participants on the role of technology in metadata in accessing oral history content. Participants were asked to express their thoughts on the technologies for metadata schema in the accessibility of oral history content. When asked to describe their experience with the metadata schema and how satisfied they were, participant C stated that: "*even though they are satisfied with metadata schema used, there is always room for improvement".* This sought participants' experiences and perspectives on the use of metadata schema for oral history content.

Participants acknowledge the importance of technology in metadata access to oral history content in their responses. Everyone agreed that metadata improves the discoverability and accessibility of oral history content. Participant D stated that: "*Technology allowed for metadata that improved the discovery of items while ensuring that the items were visible to the users."* Participant E indicated that: "*We are to link the fonds with the related ones, thus ensuring the discovery of the other items that a user might not be aware of beforehand."* Participants did not only refer to discovery and access, but also added that: "*Metadata are important to maintain consistency in the description of archival items."*

In response to the question, "*Do you think the technologies used in NFVSA effectively assist in accessing oral history content?"* This was an open question designed to ask whether technologies best suited oral history content by stating if there were gaps and/or limitations. All participants agreed that the technologies are being used effectively.

When asked to describe technologies available for use at the NFVSA, participants mentioned the information management system, digitisation technologies such scanners, metadata schema and system, and the NAAIRS.

Participants A, B, C, and E indicated that there were no technologies at the NFVSA. When probed further, the participants stated that due to system challenges they were experiencing such as regular system downtime, the lack of internet connectivity many months, and the metadata backlog, they found it difficult to keep up with creating metadata. Participant A stated that: "*We had been out of network connectivity for months now, as such the computers are there but cannot be used. Some of these computers are even obsolete and could not be compatible with the system*." Nevertheless, it was relevant to check which technologies were used to access oral history content.

4.3.5 Metadata schema framework for oral history content

The fifth objective of this study was to propose a metadata framework for oral history content in South African archival institutions. The researcher asked participants if they were aware of other metadata schemas that were used in archival institutions. Almost all participants were already familiar with Dublin Core and AtoM, as they were used at the NFVSA. The researcher needed to dig deeper to see if they had found any gaps in the existing schema, and, if so, what needed to be improved and what could be kept. Participants were asked to explain and justify the identified gaps. This was done to determine whether the participants would recommend the new standard and/or the existing schema. Participants indicated that they preferred the use of the existing metadata schema and standards at the NFVSA. Some participants suggested that the current metadata schema would be very useful, but that it should be tailored to the metadata for oral history content in archival institutions.

Given that archival staff were responsible for metadata access for oral history content, the researcher needed to know their suggestions on what would be required for the existing metadata to be effective for oral history content access. It was also critical for the researcher to ask participants if they thought knowledge and skills were needed to use the existing metadata schema and standards to ensure access to metadata in oral history. This was revealed when most of the participants, while responding, kept referring to the oral history staff for more information.

4.3.5.1 Monitoring and updates on the metadata schema

Monitoring and evaluation are critical components of the success of any activity or system. Monitoring is essential for determining what is good or bad, and it allows for keeping updated if necessary. Not all the participants specified how often monitoring and evaluation were carried out. They commented on the fact that they reported both positive and negative activities when writing monthly reports. It was clear from what was said that they did not specifically monitor the schema. "*Even though we do not have staff directly tasked with monitoring and evaluation, the policy is clear on monitoring. We do monitor but not so often because we are only doing that as our secondary job. The participant also referred to the issue of understaffing and workload."*

The question of monitoring and evaluation was less familiar and friendly to the participants, as their responses were very short, no matter how the researcher probed the subject. Most of the participants mentioned that the management was in charge of monitoring all the institution's activities, including the metadata schema. However, the participants agreed that the schema was monitored and evaluated as part of their daily metadata creation activity. According to participant G, "*So long as we work on the system on a daily basis, we are also monitoring the system because we see the updates as they happen and identify the gaps*."

4.3.5.2 Limitations of the metadata schema used

The goal was to identify which activities could not be performed by the current system. This was consistent with the study's objective of proposing a metadata framework for oral history content. Participants were asked to explain the limitations of the Atom as the NFVSA metadata schema. All participants agreed that there were some characters that were incompatible with the schema's language, which endangered the principle of provenance.

Participant A stated that even though "*It's a minor thing, but the fact that AtoM gives error messages when capturing other characters is a problem since it affects the provenance principle*". Participant A said that "*Some records are still in the Dublin Core format and not accessible. Some materials are not compatible with the new system. We still have digitisation to make materials accessible. This means that we still have a backlog in terms of content that is not accessible".*

When asked to elaborate on the issue, participant F further stated that: "*The fact that metadata are strictly done in English makes it a challenge to write some other titles that are not necessarily in English*". This was supported by Participants F and G, who added that the system lacked translation and that it would be beneficial if the translation was embedded in AtoM to help with other languages. "*This affects the creation of metadata when the metadata specialist cannot interpret and/or translate it into a comprehensible format*," mentioned participant F.

4.3.5.3 Factors in the creation of metadata for oral history content

Participants were asked to explain factors that may improve and/or hinder metadata creation and access to oral history content before proposing a metadata framework for oral history content. In general, participants complained about the complexity of oral history content for the creation of metadata. For example, participant D stated that: "*Oral history is very complex to describe. It comes in multiple subjects and disciplines in different languages with no translators or even a machine to translate, and creating metadata for oral history is just a demanding activity that takes so long a time."* The other factors are summarised as follows.

- Incompatible metadata records and oral history collections to the new system, creating a backlog.
- Most oral history records were kept in old formats, and there were no technological resources or tools to allow access to those old formats (compatibility).
- The search system did not allow linking of internal entries within or external to the system.
- Balance the creation of metadata for the older collections and the newly collected ones.
- Problems with connectivity and other system-related problems.

The results indicated that the creation of metadata at the NFVSA is influenced by a variety of factors, ranging from the metadata schema to the technicalities of the system, such as network connectivity.

4.3.5.4 Metadata schema and NFVSA policies

It was also critical to understand how policies affect metadata for oral history content. Participant E said that: "*All that we do is done according to the policies of our institution."* The question on metadata schema application aligns with NFVSA policies was intended to determine whether the policies provided a clear guideline in terms of what was applicable to

metadata access and discoverability of oral history content. Participant B stated that: "*The policy is silent about oral history content. It does not address oral history as a different kind of material, but treats it like all other audio-visual materials. To me, this is not a problem as most of these are multimedia, hence audio-visual."* Participant H disagreed *"The policies give guidelines on how to handle oral history matters. We also follow the ICA and ISO guidelines in metadata creation. The decisions taken are aligned with the policies. These include the service level agreement on the adoption of AtoM, the digitisation, etc.; in fact, everything is policy-based."*

Internal policies should guide the application of international standards in any metadata practice. However, based on the policy analysis, metadata were mentioned and were applicable to all archival materials, not just the oral history content. Secondly, this question was intended to better understand the metadata creation processes of the oral history content in relation to the institution's policies. Participant C made no mention of any policy, but stated that: *"When we identified a project, we prepared standard questions that would be used to collect information. The questions would include some metadata questions."*

All participants agreed on the importance of policy to guide the metadata schema and the metadata creation thereof. In addition to policies, participants recommended a metadata system.

4.3.5.5 Framework for the metadata schema for oral history content

The main aim of this study was to explore metadata schema for oral history content in South Africa with a view to proposing metadata schema framework. It was relevant to solicit ideas from the participants; therefore, they were asked to identify areas for improvement and make recommendations on how to improve the identified areas. It was also important to obtain participants' opinions and suggestions for modifications and/or changes that they would recommend for their organisation, the NFVSA, in order to propose a metadata framework.

Participants identified the factors associated with metadata creation and access to oral history content. The following were identified:

- There was a backlog in the metadata creation for oral history content. This happened firstly due to the complex nature and time required to create metadata for each oral history

content and, secondly, due to technology-related issues such as regular downtime and a lack of network connectivity that happened for months.

- There was also a backlog in digitisation of oral history content due to overloaded staff and insufficient skills with regard to digitisation tools. This was because there were no staff member specifically accountable for digitisation, and as such, that was an added responsibility for staff who already had their own responsibilities.
- Participants also forwarded the backlog in migration of metadata and oral history content. This was because there was much data clean-up due to system incompatibility. It was indicated that the implemented metadata schema was not fully compatible with one before it. Participant F indicated that: "Over *seven million collections had already been migrated at the time of data collection. This was a minimal number considering what needs to migrate. This was also linked to data cleaning to move from the old metadata schema, namely Dublin Core, to Atom."* This response emphasised the importance of accelerating migration and increasing staffing with the help of technological tools so that oral history content could be accessible.
- Technical loss of oral history content in the system. Some records were migrated but cannot be traced in the system and were not searchable for access.
- A lack of metadata guidelines for metadata access for oral history content
- The database search which was not user friendly, indirectly forcing users visiting the onsite for information (physical access)
- Non-extensible metadata fields
- Incompatible of the metadata schema language to the indigenous languages in South Africa

As the metadata schema used, AtoM do have a translation function, but despite that, some characters were found not to be acceptable during the metadata creation. This might affect the access to oral history content as indicated in Figure 4.4. the language of repository.

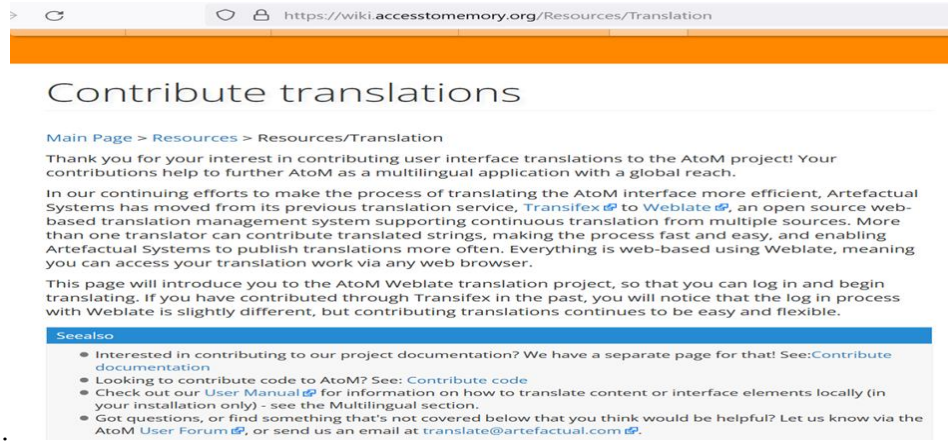


Figure 4.4: Contributions to translations

This metadata schema provides translation support for 202 languages. Except for Sepedi, 11 languages of the country are listed, while isiNdebele is divided into northern and southern Ndebele (AtoM 2023). Figure 4.4 shows the AtoM translation landing page with translation instructions.

Participants were also requested to make recommendations for the factors identified. The following were mentioned as possible solutions towards the improvement and maintenance of best practices regarding metadata schema for access to oral history content. Different participants had different opinions, but for the purpose of this study, the recommendations are summarised as follows:

- The employment of technologies to assist in creating metadata for oral history content, especially an AI, which could interpret sound and video content.
- The participants suggested that there was a need to increase the number of staff and training in the digitisation of materials.
- In addition to training in digitisation skills, participants recommended that academic institutions consider offering programmes to this effect.
- Customised metadata schema and metadata
- The participants agreed that they needed technological tools and/or software that could address the multilingual issues of oral history content, that is, technology that could interpret and transcribe the sound of the collections.

Data related to the objectives were analysed and presented. Some themes emerged from the analysis, which are discussed in the following section.

4.4 THEMES FROM DATA ANALYSIS

Thematic analysis was also used to analyse data from the interviews. According to Polit and Beck (2010:477), a thematic analysis involves "detecting patterns and regularities of themes as well as inconsistencies (done to compare similarities across cases)." To validate the results, the researcher presented exemplars of response cases in research reports. Vaismoradi and Snelgrove (2019) and Bertram & Christiansen (2017) concur on the questions that ought to guide thematic data analysis and ought to be asked continuously throughout the procedure.

- What common themes and patterns are emerging that pertain to research objectives?
- What relationship does the research's purpose have with these themes and patterns?
- Are there any instances of responses that deviate from the common themes and patterns?
- Can these inconsistencies be clarified, or can they be used to broaden or refocus the research?
- Do the themes or patterns suggest that additional data needs to be collected maybe in different field?

Figure 4.5 summarises the themes that emerged from the data analysis and presentation according to the objectives of this study.

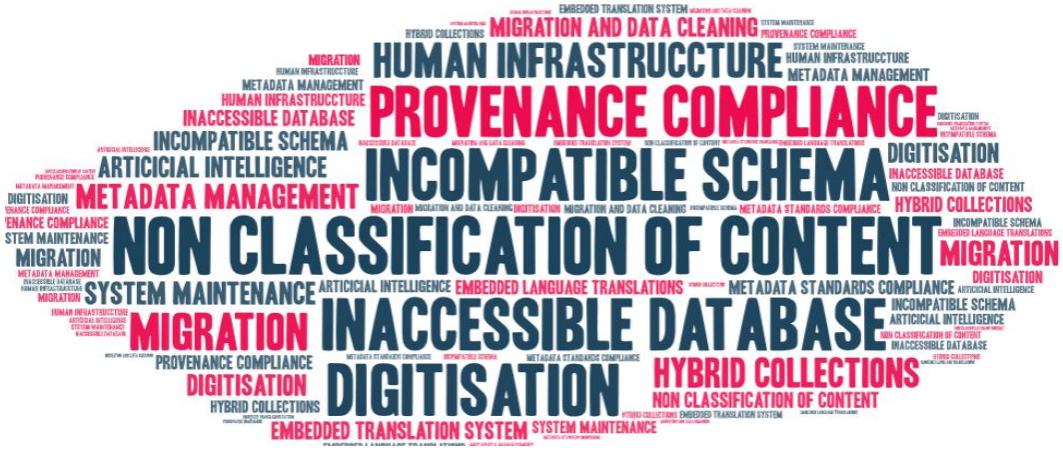


Figure 4.5: Summary of Themes (Researcher own work

The NFVSA is a public archives service available for use by all members of the public. This service to the public has to be managed efficiently, and every effort has to be made to create metadata of good quality to access oral history content.

4.5 SUMMARY

This chapter presented the results of the study based on data collected through interviews and document analysis. The data collected and presented were based on the objectives of the study. The findings of this study focused on the metadata schema for oral history content, metadata standards for metadata creation, access to oral history content, and technologies related to the metadata creation of oral history content. One of the common techniques was to identify the main themes by analysing the content of the information gathered during data collection. The next chapter focuses on the interpretation and discussion of the study findings supported by the literature review.

CHAPTER FIVE
DATA INTERPRETATION AND DISCUSSION

5.1 INTRODUCTION

The previous chapter covered data analysis and presentation, where the information gleaned from document analysis and interviews found its manifestation in line with the objectives of this study. In qualitative research, data is described through participant statements and dialogues that highlight the findings (Leedy & Ormrod 2013:313-314). Flick (2018) adds that qualitative research is not only about confirming what the literature or research has done before, but new insight should be developed of what was found in analysing data. This chapter directs its focus towards the intricate domain of data interpretation and discussion, intrinsically connected to the data analysis presented earlier.

The data interpretation process encompasses several distinct procedures: encapsulating the overarching discoveries, juxtaposing these findings with existing literature, articulating a personal perspective on the findings, and delineating the scope of limitations and future prospects (Leedy & Ormrod 2013). Therefore, the terrain of interpretation in qualitative research is diverse, adaptable to various research designs, and has the flexibility to communicate personal, research-orientated, and actionable insights.

Embedded within the realm of interpretation is the practice of acknowledging constraints within a study and charting pathways for subsequent research endeavours. These limitations are often intertwined with the methods employed in the study (e.g. insufficient sample size, recruitment challenges), emblematic of the research's vulnerabilities, which the author acknowledges to equip future studies with the foresight to circumvent similar pitfalls. The suggestions for future research serve as guiding beacons, illuminating potential research themes that could advance the existing body of knowledge, rectify the shortcomings of the present study, or pioneer novel avenues leading to practical applications or new insights.

In terms of overarching findings, the pivotal query "What lessons have been gleaned?" encapsulates the core essence of data interpretation (Silverman 2020). These lessons could come from the personal interpretation of the researcher, based on the cultural background, historical context, and personal experiences the researcher brings to the study. Alternatively, they might arise from a comparative analysis of the findings with respect to the literature or

established theories. In this way, these authors suggest that the findings may corroborate prior knowledge or veer away from it. Furthermore, these findings can uncover previously unconsidered questions emerging from the data and analysis that the researcher had not anticipated earlier in the study. This interrogative approach also finds resonance in transformative qualitative research methodologies.

Furthermore, when qualitative researchers adopt a conceptual framework, their interpretations can lead to action-orientated agendas for reform and change. Researchers could elucidate how narrative results align with the prevailing theories and the larger corpus of literature on the subject. In numerous qualitative paradigms, researchers also engage in an insightful discourse on the literature, concluding their study with such discussions (as expounded in chapter two).

5.2 METADATA SCHEMAS USED FOR ORAL HISTORY AT THE NFVSA

The metadata schema was investigated because the researcher had to describe the elements used for metadata creation and the categories of metadata applicable to oral history content at the NFVSA. The use of a metadata schema goes through "various standardisation phases, from which more or less binding metadata standards developed" (Bender, Hausstein, Hirsch& Bundesbank 2018). Metadata schemas vary in the type and quantity of data elements, the fields, encoding specifications, and the application of data content and value standards (Miller 2022; Xie & Matusiak 2016). Additionally, the unique nature of oral history content and the different ways in which it can be organised require a metadata schema that would allow possible methods for restructuring of the descriptions, thus enabling discoverability of, and access to, these collections (Fell & Ou 2022:227). Therefore, a decision to adopt a particular metadata schema has implications for the metadata quality and level of description. Palavitsinis Manouselis and Salvador (2017) explain the importance of maintaining the quality of metadata in maintaining quality metadata repositories. Hence, it was important for this study to describe metadata schema in use at the NFVSA so as to determine whether the schema adopted needed to be customised. Furthermore, the description and decision to see whether the metadata schema is a best suit for oral history content, familiarity with the schema is required. Thus, this study collected data to determine whether the staff at NFVSA were knowledgeable about the metadata schema in use.

5.2.1 Knowledge of metadata schema

The research results showed that despite the fact that the metadata schema in use was still in the implementation stage, all participants were familiar with the schema in use. On the other hand, even though participants held qualifications, some were not in archival education. This was not surprising, as Maluleka, Nkwe and Ngoepe (2018) established that the education programme available for archivists in South Africa consisted of only a few programmes offered by four out of 26 institutions of higher learning. Only one of the four South African universities offering archives and records management qualifications, provides a full bachelor's degree with a major in archives and records management (Ngoepe & Saurombe 2021). This is contrary to Zimbabwe where institutions of higher learning, including polytechnics and universities, offer records and archives training programmes and churn out hundreds of records and archives professionals each year (Tsvuura 2022). Despite the participants having post-matric qualifications, they relied on short courses or on-the-job training and workshop to enhance their knowledge of archival systems. This could be that employees of the archival institutions organization were regarded as professionals if they had completed internal training and had a Bachelor of Arts (BA) degree with a history concentration (Katuu 2023). Additionally, Ngoepe and Katuu (2017) identified lack of ICT skills from archival staff to function with digital technologies.

Metadata schema are web based thus requiring knowledge and skills aligned with the digital environment. Due to lack of educational basis on metadata schema and oral history content, according to Ngoepe and Katuu (2017) archivists and records were not consulted when ICT systems are installed in institutions due to their lack of knowledge regarding technology applications and how best to incorporate them into recording keeping activities. Furthermore, it is frequently the case that records administrators and archivists are inadequately able to converse with ICT specialists in concepts and language that both parties can understand (Ngoepe& Katuu 2017). Hence, the researcher deemed it valuable that those staff would be required to attend workshops on the metadata scheme since its implementation, and the ICT relating to its functionality.

Regarding work experience, participants had a considerable amount of experience, as most had been employees before and after the new metadata schema was rolled out. This suggests they had sufficient experience to apply their knowledge gained from the previous schema to learn

the current one. The participants gained the knowledge from interacting with and using the schema, or by training. As already stated, metadata schema can assist to improve the quality, consistency, and usability of the metadata. It can also facilitate collaboration and interoperability with other systems and/or platforms (Knight, Rodrigues & Ciota 2020).The challenge remains that if metadata practitioners lack knowledge on the metadata schema and/or oral histories, this will result in poor metadata thus affecting access.

Another lack of knowledge relates to awareness on the available and/or existing oral history content at NFVSA that are uncatalogued. Given that some oral history collections were never described and//or classified, they remained in boxes and as such some staff did not know the oral history content available in those boxes. Without the descriptions and/or metadata, oral history collection remain inaccessible. This is in addition to the already existing of metadata gaps which affect metadata quality contributing to lack of access to oral history content. Poor metadata impact negatively on access to archival collections (Mazé 2014). This will also overthrow archival institutions of making archives accessible. Therefore, it was necessary to describe the metadata schema used at the NFVSA.

5.2.2 Metadata schema used at the NFVSA

Metadata records are created after a metadata schema has been selected, customised, and documented (Xie & Matusiak 2016). Hence, it was necessary to describe the metadata schema in use at the NFVSA. The AtoM schema was adopted at NARSSA, and since the NFVSA is a subsidiary of NARSA, it was important for the NFVSA to also implement it. Participants agreed that the metadata schema in use was AtoM even though it was not fully rolled out at the NFVSA due to a network issue. This delay has led to some archival institutions being reliant on Dublin Core, which was phased out, as it did not cover some fields that were necessary for quality metadata. The transition from Dublin Core was not without challenges. Some data clean-up was done, as the two schemas were not compatible.

Even though some staff were never fully active on the metadata schema system, their description of the current metadata schema reflected their knowledge on the previously used schema. The current metadata schema seem not to fully cater oral history content. This was why metadata elements used for oral history content had to be identified.

5.2.3 Metadata elements for oral history content

Metadata conform to some set of specific metadata in a metadata schema. Many of the metadata elements had useful information for a range of situations. At the NFVSA, all elements of AtoM are applicable to the metadata of oral history content. At the time of data collection, some participants indicated that they had never created or captured new metadata records due to poor network connectivity but acknowledged that all elements in AtoM are relevant to the description of oral history. The question as to what happened to the oral histories that were collected arises. This would create a backlog in metadata, which would negatively affect access to archival collections. A well-compiled metadata description for various oral history content is a requirement for access to any resource, including oral history content (Herawan & Sensuse 2018). Demichev, Kryukov and Prikhodko (2019) state that there is sometimes it can be difficult to identify the essential elements required to fully describe a resource. Such identified elements should describe the origin, composition, and navigation, as well as the digital object's quality, longevity, validity, and recognition. On the other hand, AtoM is more "complicated and requires more training and knowledge, but it gives more extensive and detailed metadata," which is specifically useful for describing archival collections (ICA 2016). This could suggest that archivists might not have completed other elements, as the NFVSA was still at the AtoM implementation stage and/or due to data clean-up required after the migration from Dublin Core.

5.2.4 Metadata creation process

The process of metadata creation process is summarised as: determining resource characteristics; transcribing available descriptive information; conducting subject analysis; selecting appropriate terms from a designated controlled vocabulary tool; following content guidelines for data entry; recording administrative and preservation information; and adhering to the established standards. This study found that the metadata process and the NFVSA followed the said steps as indicated by Xie and Matusiak (2016).

5.2.5 Metadata categories recorded

Descriptive information is used to find things, structural metadata is mostly used for storing and presenting objects in a repository, and administrative metadata is used for administering and maintaining objects in the repository (Xie & Matusiak 2016). According to Riley (2017),

libraries and archives have a long history of creating and sharing robust structured metadata. To access resources efficiently and accurately, it is important that structured data are used to describe the resource.

At the time of data collection, the NFVSA had not yet created ant records even though Atom was officially launched, as they experienced continuous system downtime and network challenges. However, the metadata records created on the old metadata schema were migrated into the new one. This was also clarified on the NARSA website, which explained information fields between the old system and the new one.

The conceptual study was done to determine types of metadata and identify the core metadata elements used to describe oral history content at the NFVSA. The results showed that the NFVSA included administrative, descriptive, and structural metadata in its description. Unless the content is documented with descriptive, structural, and technical (or administrative) information, a digital object has no significance for a human person (Berry 2018). As archival institutions should create metadata to organise and manage it globally, metadata will always remain a crucial component. There is no single type of metadata that is applicable to every applications, knowledge communities, and information source.

Hence, even when staff were not familiar with the schema, they were able to describe the elements based on their knowledge of the old metadata schema. The metadata schema has embedded metadata standards for metadata creation purposes. Therefore, the next section is a discussion of the findings relating to metadata standards used for the classification of oral history content.

5.3 METADATA STANDARDS FOR ORAL HISTORY CONTENT AT THE NFVSA

This study revealed that oral history content was not classified at the NFVSA. Metadata standards for classification purposes were not created with classification in mind. The classification was done through titling of the collections. Keywords and/or search terms were derived for the title of the content at hand. No external search terms were added because of the principle of provenance and principle of originality.

5.3.1 Knowledge of metadata standards

Participants showed some knowledge of the metadata, and thus training was required to familiarise the staff with the system. Ahmed (2018) agrees that a lack of knowledge of archival standards affects metadata classification of archival collections. Participants seemed not to be have had differing ideas on what classification is, and the purpose thereof. This was not surprising, as Ngoepe (2019) indicates that classification of items is often misunderstood, leading to oral history content to remain unclassified. It was clear from this study that classification of oral history content was not given most preference.

5.3.2 Classification standards for oral history content

The findings showed that the NFVSA used the standards that were supported and embedded in AtoM, namely ISAD(G) and ISDIAH. While archival resources are valued for their uniqueness, they do not exist in isolation and stand to benefit from linked data treatments capable of exposing them to a larger network of resources and potential users (Dobreski et al. 2020). Standards should describe and classify these materials to ensure their discoverability and accessibility. As a result, it was appropriate for this study to ask participants about the application and role of these metadata standards in the accessibility and retrievability of oral history content.

Keywords were seen to be classification terms despite not referring to the content but to the description of the oral history content, that is more of the administrative metadata type. There are five categories based on common uses of metadata to facilitate and increase interoperability: accuracy, currency, provenance, privacy, and classification (Király 2019). Each category of metadata elements is critical for describing the attributes across all sources to ensure interoperability (Simons 2014). This was not the case at the NFVSA. Provenance dominated the classification and use of keywords outside the document itself. The staff argued that due the principle of provenance they could not use any term that was not used from the original source. Until provenance is addressed in relation to accessibility, then that would remain a problem.

The objective of this study was to analyse the metadata standard used at the NFVSA for oral history content.

5.3.3 Classification guidelines for oral history content

It was previously understood that the NFVSA had a file plan to classify oral history content. The functions of various government departments and/or organisations are used to develop file plans. If file plans are designed to fit the functions of a department, it is a concern whether oral history content is correctly placed in archival institutions in terms of accessibility.

File plan is official guideline for classification in archival institutions and has been officially approved by the National archives. Unfortunately, it is not the best fit for oral history collection. This suggests that archival classification will remain unachievable for some time and so is description and access oral history content. The findings are in line with Ncala (2017), Ngoepe (2019), and Ngoasheng, Ngoepe, and Marutha (2021) which found out that archive materials were inaccessible because they remained unclassified. A study conducted by It is thus relevant for this study to propose a framework to address classification challenges of oral history content. Given the findings of this study, one could argue that the NFVSA was responsible for the collection of oral history content, but the adopted metadata were for archival materials in general and did not take into account the nature of oral history content.

5.3.4 Choice of metadata standards

Archival image collections are particularly challenging because very few items have individual annotations, and the level of consistency and accuracy of description may vary from item to item. Finally, metadata records need to be reviewed for quality and consistency. Ultimately, the quality of metadata and adherence to standards determine if digital objects are findable and discoverable within local digital collections, as well as in the aggregated environment of large-scale digital libraries.

5.4 ACCESS TO ORAL HISTORY CONTENT AT THE NFVSA

The conceptual framework was used to determine the classification standards used and whether the metadata principles, such as interoperability, refinement, and extensibility, were followed. Knight (2022: 116) states that: "*where the physicality of the describing record is also removed, access can be enhanced by searching for any of the facets or data elements used to describe the information object*." This suggests that subject organisation and/or classification become crucial tools. Therefore, Knight (2022) adds that while keyword searches can be executed on different facets or combinations of facets in a database, uncontrolled keyword searches will

often retrieve results referring to unwanted and irrelevant things. This is because even although the NFVSA does include keyword searches as uncontrolled vocabulary, they do not produce the intended result. Oral history content and collections are part of archival collections, which are among the most valuable materials held by archival institutions. As indicated earlier, the oral history content is, by nature, unique and possibly of high importance in terms of academic, research, financial, and/or social value.

5.4.1 Formats of oral history content

Oral history content in the NFVSA was in both analogue and digital format. Some of these oral history contents were still available on audiocassette tapes, minidiscs, and mini-DVDs. This suggests oral history might not be accessible or discoverable by the public, as they might not be aware of its availability and/or existence. The researcher confirmed the formats of the oral history content and found that some were still contained in older sources such as audiocassettes and, as such, not compatible with current tools for access. Furthermore, those sources did not have metadata records and were therefore not accessible when searching the database.

5.4.2 Search strategies to access oral history content

Access to oral history content is critical to identifying, recognising, and respecting various aspects of heritage (SAHO 2019). One of the primary services provided by national archives is access to archival material. Although oral history content is intended for researchers to use, many institutions, including NFVSA, struggle to provide access. Inadequate access is caused by a variety of factors, the most significant of which is a lack of agreement or standardisation in classification methods. Metadata facilitates access to oral history content because the descriptions are searchable from any database. Formats were also a contributing factor in accessing oral history content.

5.4.3 Ethics and legal access to oral history content

The extent of the collection is immense, containing an array of collections of archival sources pertaining to, especially, the history and development of Cape Town and the Western Cape province. In addition to official government records, substantial source information is also available for genealogical or family research. The NFVSA, as the guardian of oral history collections in South Africa, should ensure that it complies with its responsibility of safeguarding these collections and making them more accessible for current and future

generations of the South African state and society (Hendrich 2017). It is for these reasons that the NFVSA should address the accessibility and searchability of the database. Ngulube et al. (2017) state that archival institutions manage archives to be used although they face challenges such as giving access to their holdings and reaching diverse users. The lack of finding aids added more challenges to accessing oral history. Some oral history content was not described; hence, for some, there were no inventories.

In addition, some oral history content remained inaccessible because the information contained was sensitive. On the other hand, access to oral history content was hampered by the absence of legislative framework detailing the exact permission, ethical procedures, and requirements available to the users. Participants agreed that the existing records management policy and the NARSSA Act did not hold direct accountability for oral history access. Access to, and creation of, metadata for oral history content could be explained in terms of technology, as it relates to database and information systems, as well as access to oral history. As a result, the following session focuses on data collected on technologies for accessing and creating metadata for oral history content.

5.5 TECHNOLOGIES USED FOR METADATA SCHEMA OF ORAL HISTORY CONTENT AT THE NFVSA

Metadata standards, whether for the description, management, or preservation of digital objects, are critical technological resources in interoperability (Formenton & De Souza Gracioso 2022). The NAAIRS of South Africa is governed by section 3(e) of the NARSSA Act, as amended.

Compared to other countries, such as Zimbabwe, the legislation is silent on automating the detection of aids. Automating the finding aids can help off-site researchers determine whether to visit the National Archives. Meanwhile, NARSSA makes both databases available to the public. Using the “Search Old NAAIRS database” will provide a more comprehensive search, whereas using the “Search New database” will allow the public and participating institutions to provide feedback to NARSSA regarding potential problems that may have occurred during migration. The issue remains that some of the new database’s items do not display all the records in the system. Accurate and complete metadata are central to migration; thus, metadata describe the attributes of the materials from which one can derive the type of system in which it can run (e.g., the operating system), the versions of any software dependencies, and other

criteria that are crucial for accurate creation of in the digital environment (Hart & De Vries 2017). Hart and De Vries (2017) further indicate that metadata records can easily be lost and corrupted by everyday use.

5.5.1 Technologies for access to oral history content

Archival institutions use many types of technologies. Understanding the technologies used for metadata around providing access to oral history content was also important, based on the participants' experiences. Literature demonstrates that technology had a significant impact on how archival institutions provide their services and functions. The conceptual framework served as a foundation for identifying the technological tools used at the NFVSA for the metadata schema for oral history content. It became clear that technologies were used for the description of metadata content and the creation of metadata for archival materials, including oral history content. The metadata schema used at the NFVSA was standards-based description and access. The online search was available, but was affected by challenges such as system downtime. The fact that more than 8.3 million entries still had to be migrated from the old NAAIRS database to the new database also affected access to oral history content. It was also found that migration was very slow because of technologies and the incompatible nature of the metadata system. There was a need for digitisation for which the technologies were available, but there were very few staff members with knowledge to use these technologies. Technological tools available at NARSSA were the information management system, digitisation technologies such as scanners, metadata schema and system, and the NAAIRS. This was because the NFVSA depended on NARSSA for its information management system. Hence, it appeared that the NFVSA had interacted with the system less frequently. It was still possible to analyse the metadata records for the country as the metadata from provinces were deposited at the national archive.

5.5.2 Finding aids and access to oral history content

The objective of the study was to determine access to oral history content at the NFVSA. Finding aids are used to search for information in archival materials, but despite the availability of the search function, some oral history content does not have inventories. The finding aids are not searchable or are difficult to search from; therefore, it could not assist in access to oral history content. It was also important to determine the search strategies that users could access while searching for oral history content**.** Access to oral history content requires effective search

strategies and searches. Most importantly, the retrieval of records in the NAAIRS is influenced by search terms. Depending on the scope of the information required, users may use single or multiple words. The system included search strategies for Boolean operators. Users were unable to access the oral history content with their formulated search words because metadata standards did not use standardised lists of words. Given this, participants were asked about the process of adding search terms to the NAAIRS. The principle of provenance remained important, which affected the creation of new terms that were originally not on the content. Terms were only derived from the title of the archival group. It is for that reason that no subject entries for oral history content and/or archival materials were created. Additionally, the database was not user-friendly and users did not know how to navigate through it. Molefe and Schellnack-Kelly (2019) concur that NAAIRS and AtoM as online finding aids were considered not user-friendly. Some participants agreed that the database was not user-friendly and further indicated that users were forced to request services from the archivists by physically visiting the institution.

5.5.3 Technologies for metadata creation

Metadata creation was processed online through an open-based metadata schema. It was easier to keep the metadata updated, but the complex nature of metadata creation for oral history content affected the current status of metadata records. Technologies were necessary to ensure updated and accessible metadata records for easy access of oral history content. The NFVSA has digitisation tools for improvement of metadata creation and access. These benefits may be attributed to the use of technologies for oral history content.

The important role of technology in metadata creation and access to oral history content cannot be overemphasised. Databases improve the discoverability and accessibility of oral history content.

5.5.4 Benefits of technologies used for oral history content

Rapid digitisation at the NFVSA was necessary. The lack of usability and external accessibility of the database were barriers to accessing not only oral history, but also the national archives in general. Users must physically visit the site, which has an idirect impact on their time and expenses. As a result, Mosako and Ngoepe (2021) focused their survey on how the public could remotely access archival material. Mosako and Ngoepe (2021) emphasise that the adoption and

use of digital platforms allow continuous open access to archives from anywhere in the world with minimal commuting and physical human interactions, and institutions must reinvent themselves for 21st century audiences. Mukwevho and Ngoepe (2019) found out some archival institutions were interacting with users on Facebook, LinkedIn and Twitter even though relying on social media managed by the mother bodies without own social media pages.

Another barrier to access was item compatibility with the system, and files can sometimes be too big to be accessible on the system. Cloud storage helps to improve access to cloud resources (Shibambu & Ngoepe 2020). Ryan, Keating and Finegan examined the creation of a selective web archive in the Irish web space. The study discovered that the unusual legal context in Ireland in terms of legislative barriers to archiving the Irish web space at scale had significant implications for data loss and long-term access to Ireland's contemporary record. The need of the NFVSA to use cloud storage as the new method of accessing and interacting with archival collections on and off site, necessitates policy development and amendments. It is clear that the NFVSA, through NARSSA, needs to investigate the possibility of producing materials remotely. This should be considered within the ethical constraints that exist on the system of some of these collections. It can be suggested that archival institutions should consider embedding ethical compliance processes to be conducted online.

The use of technologies brought about change in access to oral history content. The results of this study showed that unlike other archival materials, there were no inventories for oral history content. Without the compatible hardware and software, as well as digitised oral history content, access remained a challenge. Some formats were no longer accessible, as they were not compatible with the new technologies in our organisation. This had a negative impact on access, as some of the collections were either hidden in unprocessed collections and may not be available for years because the bigger the backlog, the more technology changes (Ngoepe 2019). This leads to inaccessible archival knowledge and sources. Borrowing from the study by Mosako (2022) focused access to heritage knowledge access in museums, this author recommended the implementation of ICT to provide educational services to communities. Furthermore, established "online communities can enhance intellectual growth," (Alfier & Feliciati 2017), thus ensuring open access for communities in digital technological environment (Opgenhaffen 2020). The researcher is of the opinion that even within archival institutions oral history knowledge may be made available through the use of technologies.

5.6 PROPOSED METADATA SCHEMA FRAMEWORK FOR ORAL HISTORY CONTENT

The findings indicated that there was a need for review of the metadata schema and standards to ensure improved access to oral history content. Various factors such a backlog in metadata creation, digitisation, and migration were mentioned as affecting access to oral history. In addition, understaffed and non-skilled staff as well as the language barrier may negatively affect access. For example, some collections could not be accessed without the assistance of a linguist, and very few archival staff were knowledgeable about the languages in which the records exist because most users were not familiar with Afrikaans and would like to have the descriptions of the archivalia duplicated in English so that they could assess the contents of the material (Molefe & Schellnack-Kelly 2019:12).

The existing metadata schema was preferred by the staff at the NFVSA despite factors relating to its implementation. The implementation of the schema faced challenges such as:

- knowledge and skills of the metadata schema and standards
- knowledge and skills of oral history content
- staffing and workload
- continuous monitoring and evaluation of metadata schema
- language compatibility as some characters were incompatible with the schema
- principle of provenance.

Additionally, the metadata creation was affected by the complexity of oral history content. The multiple subjects and disciplines in different languages played a deterring role in metadata creation and access to oral history content. The search system did not allow linking of internal entries within or external to the system and classification of the content. There were also challenges of complicated database search, network connectivity, and system unavailability for metadata creation, access, and discoverability. For example, Mosako (2023:201) emphasises "the importance of interconnectivity and states that online interconnectivity serves as a crucial link between cultural institutions and society, as it is via such connectedness that individuals engage with and study culture and history, which serves as a template for shaping society's awareness and behaviour." This is relevant even to archival institutions and the NFVSA. The network connectivity

had a serious impact on the use of the metadata schema, resulting in the disconnect between the NFVSA and their users, as no metadata were created and/or updated.

On the other hand, the issues of policies that were not specific to oral history content remained a problem. Oral history content is hidden under the records at archival institutions, which led to a lack of guidelines in handling metadata creation for oral history content while ensuring metadata schema that would be compatible with the unique nature of the said content.

Internal policies should be revised to guide the metadata schema and the creation thereof. Hence, the aim of this study was to propose a metadata schema framework for oral history content in South Africa

Factors aimed at developing the metadata schema framework for oral history related to:

- the need to address the backlog in the metadata creation for oral history content
- strategies to address the backlog in the migration of metadata and oral history content
- managing technical issues such as system downtime and a lack of network connectivity
- increased staff and accelerated digitisation of oral history
- upskilling of staff with regard to digitisation tools
- AI for reduced time and effort on metadata creation for oral history content
- increased technologies for metadata creation and access
- expansion of physical access to digital access to archival institutions
- language and translations AI assisted
- customised metadata schema and metadata standards for oral history content.

The above issues regarding metadata schema and standards were elaborated as follows:

(a) Digitisation at the NFVSA

There have been reports of digitisation delays due to a variety of factors. The migration of metadata from Dublin Core to AtoM resulted in a significant amount of data cleaning. As a result, both the old and new databases ran concurrently. Because some of the materials were inaccessible on the new system, digitisation was considered necessary. The archives had to search for and download materials, which would then be uploaded to the NARSSA public website for easy access by users.

(b) Multilingualism and translations challenges

Data showed that participants were concerned that the Metadata schema would refuse to accept certain other characters, which would violate the principle of provenance. “Language is an essential component of any worldview,” writes Knight (2022:106).

Clients of DataCite were strongly encouraged to provide metadata in English whenever possible, in addition to any other language that the funder or hosting organisation may require (DataCite 2021). The question is whether the language experts will be available and accessible for this purpose. The oral history content collection is diverse, with objects in various formats and languages, as well as descriptions formed by various metadata practices. The metadata can be used to identify alternative versions of resources in different languages, as well as the source of the translations. Translating relevant specifications and standard documents into a variety of languages is a good place to start when promoting a global metadata architecture. The Dublin Core Metadata Initiative (DCMI) keeps a list of basic document translations. Similarly, the European Workshop on Learning Technologies maintains translations of the Learning Objective Metadata (LOM) specification. Metadata, for instance, can specify the resource's language and character set. Alternative language versions of the resources as well as the source of the translations may be identified by the metadata. It is thus relevant to capacitate staff in various language or language translations. This may be a laborious activity the require additional and trained staff.

(c) Staffing: understaffed, training, workload, and backlog

The human element is a requirement in digital preservation, as certain metadata, such as descriptive and administrative metadata, can only be created by humans (Hart & De Vries 2017).

Since descriptive and administrative metadata can only be created by humans, the human element is required in metadata. Oral history metadata creation becomes difficult due to the difficulty of describing those interactive elements unique to oral history content for which no ISAD(G) element exists. Chapman (2021), for example, emphasises the importance of recording a descriptive summary, or “purpose or intent of use”, whereas Robson, Sassen, Thomale & Yanowski (2019:208) emphasise the often overlooked importance of “duration of play, number of players, and age of players”. These characteristics can help improve record

retrieval and access (Chapman 2021). Making oral history content more accessible involves more than just selecting, digitising, describing, curating, and making their holdings available online. It also includes balancing these activities with the ongoing acquisition and processing of large backlogs of traditional materials, as well as figuring out how to identify, acquire, preserve, describe, and disseminate born-digital materials of archival value (DataCite 2021). The findings pointed to staffing as one of the causes of the backlog in metadata creation and digitisation. Training in the metadata schema, digitisation, oral history content, and legal framework needs to receive more attention in archival institutions, especially toward oral history content stored on analogue formats.

(d) Digital transformation: AI for reading and interpreting audio and long text

Oral history description can be time-consuming, especially when a detailed and accurate recording of the fonds/materials of the objects, object numbers, colours, measurements, and materials, and an accurate record the content type, size, materials, and extent of an item are required (Király 2020). Furthermore, the findings raised awareness and emphasised the necessity for caution and expertise when handling digital data by highlighting the importance of provenance metadata.

(e) Legislation and metadata and access to oral history

Archivists must keep up with the increasingly complex and multi-national policy issues surrounding ownership, provenance, privacy, and cultural sensitivities (Knight 2022). Metadata and access should be consistent with the institution's policies. The findings indicated that the NFVSA should be take note of policy awareness. If participants were unfamiliar with these policies, it could have an impact on compliance. This was also consistent with understanding of the metadata standards used in metadata creation. The study argues that it is not only the communities who need to be educated about the laws protecting their oral knowledge, but also the archivists (Zimu-Biyela 2022). In general, the users may not be aware of the available oral history in South Africa as their metadata records remain incomplete and/or non-existing. Even at the archives, oral history content remains marginalised and does not receive special attention due to its nature. Attention is mostly focused on other archival records, hence the file plan was designed to fit the record management purpose. Halamy and Kibat (2017) describe oral history being of no commercial value as the reason for limited effort from

archival institutions and the community in general. Halamy and Kibat (2017) recommend that oral history should be included as a subject in schools, either secondary or primary schools, to raise awareness of the importance of oral history among these young generations. The researcher believes that, in addition, metadata for oral history content should also be given more attention in institutions of higher learning.

Generally, most of the participants agreed that there should be training on metadata for oral history content and expressed regret that most training for oral history content is done outside of the archives, despite the fact that it is part of the archival collections. It was also stated that digitisation knowledge is important and that all employees should have such knowledge. A follow-up question was asked about the ability to digitise oral history materials. Most oral history collections are organised by topic and/or region, which implies that existing metadata standards would need to be modified to accommodate the unique nature of oral history content.

5.7 SUMMARY

This chapter interpreted and discussed the study findings as presented in chapter four in order to provide readers with a clear understanding of the data. The data discussed as part of the study's objectives included the metadata schema and standards used for oral history content, access to oral history content, and technologies used for metadata schema of oral history content. In the following chapter, a proposed framework for the metadata schema for oral history content is presented after the summary and conclusions of the findings. The interpretations and discussions of the research findings of this study are presented in chapter five. Based on the literature reviewed in chapter two and the conceptual framework for this study, the data presented in chapter four are discussed and interpreted. According to the researcher, despite metadata schema and standards, there is still limited access to oral history content. This could refer to incomplete metadata records, oral history formats, or poor or non-content. The poor level of classification guidelines by metadata standards due to a backlog due to staff capacity, system downtime, and network connectivity added to the difficulties. There was also delayed digitisation, which added to the existing problem of inaccessibility and discoverability. The level of community ignorance about the role of archival institutions will persist if the NFVSA remains an institution that only researchers are familiar with. Although most respondents indicated that they used some kind of descriptive standard, oral history is described using different approaches, at different levels, and at different stages of the oral oral

history metadata creation. Although metadata are used to support the preservation, access, and use of archival materials in the future, the history and provenance, and the search and retrieval functionality, this study focused on metadata being used to support access to oral history content. However, access to oral history content is difficult if adequate metadata are not present, especially if the item is outdated and no longer supported. Looking towards the future, the NFVSA should ensure that metadata are correctly processed for oral history collection.

The following chapter focuses on the summary, conclusions, and recommendations for this study.

CHAPTER SIX
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

6.1 INTRODUCTION

Data are presented as evidence for the conclusion to be drawn (Leedy & Ormrod 2013) The findings presented in chapter five were discussed based on the data analysis and presentation in chapter four, as well as the literature review in chapter two. According to Tracy (2019), all loose threads are gathered because research must eventually return to its beginning point. The current chapter is organised as follows: a summary of the findings, conclusions based on the findings of the study, recommendations, proposed framework, and suggestions for future research. Furthermore, this chapter proposes a framework that it is hoped will become an enabler for archives to adopt a metadata approach with the intention of making oral history content accessible without challenges of incompatible format and misplaced content that is not linked to the items described, as well as well accessed at their convenient location and space. The chapter also makes recommendations for future research based on the findings of this study and discusses the practical implications of the findings and metadata access in archives.

This study was aimed at exploring the metadata schema for oral history content in South Africa, with the aim of proposing a framework to ensure its accessibility and retrievability. To achieve the aim of this study there are research objectives that served as a guidance for this study. These specific objectives are to:

- describe the metadata schema used for oral history at the NFVSA
- analyse metadata classification standards for oral history content at the NFVSA
- determine access to oral history content at the NFVSA
- determine the technologies used for metadata of oral history content at the NFVSA
- propose a metadata schema framework for oral history content.

6.2 SUMMARY OF THE FINDINGS

The researcher provides a summary of the findings in accordance with the objectives of the study.

6.2.1 Metadata schemas used for oral history content

The main objective behind the creation of metadata records and context is to improve information access. The first objective of this study was to "describe the metadata schemas used for oral history at the NFVSA." The intention was to establish the metadata elements that are applicable to the oral history content and the metadata categories. Document analysis and interviews were used to collect data to ensure that the research objective with respect to the metadata schema for oral history was addressed in detail. The aim was to identify the fundamental elements needed to describe oral history content. Most of the participants indicated that all elements are applicable to the metadata of oral history content. What did not come out clearly was whether the current metadata schema is perfectly suitable for oral history content. This means that the metadata schema used did not fully fit the metadata requirements of oral history. Despite having changed the metadata schema, the NFVSA did not successfully implement the new schema, as they were always encountering network challenges and system downtime. As such, they could not confidently agree that the metadata schema was the best fit for oral history collections. Linked to the recommendation provided in this study, it could be deduced that the metadata schema used needed to be reviewed for adaptations and improvement.

6.2.2 Metadata standards for oral history content

Metadata standards are useful tools to raise awareness of oral history collections. However, the available standards are geared towards archival collections in general. Metadata standards, on the other hand, are intended to guide which information to include and how to include it in each metadata element. It is obvious that metadata standards cannot include information that is not within the scope of the metadata schema. Metadata standards are also embedded in the metadata schema. As a result, whether they fit or not, archival institutions will use the standards. This explains why some employees used the standards but were not aware that they existed. It can also be concluded that the metadata used do not fit the metadata creation for oral history because they were used with schemas, which did not fit oral history collections. It was found that the standards in use were ISAD(G) and ISAAR, and that they did not aid in the classification of oral history content. As a result, such collections would be unknown to the communities. This is consistent with Schellnack-Kelly (2017:113) stating that in South Africa,

there is concern that awareness and respect for the country's cultural diversity has dwindled, necessitating the need for collaborative partnerships to address these gaps.

6.2.3 Access to oral history content

Access to oral history content is critical for citizens to understand their origins and identify them while appreciating different cultures and heritage in a country. Heritage is defined as a "consciousness of the past that gives us individual and collective identity," allowing us to reflect on what the past means to us and our national identity in various ways (OHA 2019). According to Ngoepe (2019:152), oral sources are an important part of "recreation and rethinking of the past, particularly for those who never had the opportunity to tell their stories in the past." Oral history access is critical to preventing further loss of African knowledge and many narratives (Ngoepe 2020). Oral history content can be easily accessed through public archives. "Access to primary data contained in archives as constituted in documents, housed in archival buildings, and managed by a necessary archival institution is an important component of archives administration" (Ngulube, Sibanda & Makoni 2013:124).

Despite South African efforts to ensure that oral histories are collected, access to the collection remains a challenge. The reasons for the oral history content not being accessible are threefold: (i) there are no metadata records for these collections or the metadata records do not provide enough information to make them discoverable; (ii) some of the collections are still in the old format and are incompatible with the current access tools (analogue versus digital); and, finally, (iii) some of the oral history content collected cannot be traced back to the NFVSA.

On the other hand, oral history content is hidden beneath records and other archival materials in South African archival institutions. This is because, despite its unique nature, the metadata creation process for oral history content is the same as for other archival materials. As a result, the metadata held at the NFVSA does not fully describe oral history content and is not easily discoverable oral history content. Furthermore, users do not have easy access to the ethical and legal processes that must be followed when accessing oral history. Although the search functions are unfamiliar to a novice user, archival institutions appear to be geared towards researchers. This is related to the NAAIRS, which is not user-friendly and requires users to visit the site to gain access. In summary, oral history content remains inaccessible due to a lack of inventories, where available, and the collection's delayed migration and digitisation. This is because some oral histories are still in old formats that are incompatible with the current

system. Woods, Chassanoff and Lee (2013) indicate that many of the challenges often arise from working with implementations of metadata schemas that have been designed primarily to accommodate the acquisition, analysis, and transformation of individual files. In addition, oral history content can contain several files with many potential internal dependencies between documents. On the other hand, metadata standards used did not allow subject access through content classification or keyword identification, and oral history is not browsable. This is because oral history content does not fit easily in the existing metadata schema as these schema and standards were developed for other types of materials. However, an archival institution should, where possible, include indexes with time tags relating to the recording, thorough descriptions of the content, or other guides to the contents in order to improve accessibility and retrievability of the audio or audio/video files. This access to oral history is also affected by NFVSA technologies. This will be discussed in more detail in the following section.

6.2.4 Technologies used for oral history content

Technology has changed the way archival institutions operate and provide services. The metadata for oral histories was initially displayed manually. Access to metadata has evolved over time from different formats to digitally accessible formats. This evolution is not unique to the NFVSA, from manual to Dublin Core, and then to AtoM, an open-source software. The AtoM system is a metadata encoding standard for capturing descriptive information about archival materials, such as oral history collections. As a result, such technologies were found to be ineffective in terms of metadata creation and access to oral history.

According to the results, all metadata records were successfully migrated to the new NAAIRS, the AtoM-supported database. However, it started with a system crash and, when that was resolved, data clean-up was the only option. There were more than seven million migrations of archival groups/fonds, which participants stated was equivalent to a 6 km run in a 100 km race. This was linked to the lag in digitisation. Some of the oral history content cannot be accessed from the database because the database was not user-friendly, and the oral history content that can be accessed through the metadata may not be accessible to do so due to old format challenges and incompatibility with the new system. These always require the user to physically visit the archival site rather than accessing it remotely. According to Mosako and Ngoepe (2021), archives should consider adopting and using digital platforms to provide access to archives from anywhere in the world with minimal commuting and physical human

interactions, and institutions must reinvent themselves for 21st century audiences. This is especially true since NFVSA users were still required to physically visit the archive and NFVSA technologies did not help to access oral history content.

This study also discovered a backlog in the creation of metadata and the digitisation of oral history content. Migration from Dublin Core to AtoM caused the backlog. The change from the previous schema required extensive data cleaning because the two were not the same. AtoM required more metadata elements than Dublin Core. The backlog of metadata was also mentioned in relation to NFVSA staffing. It was found that staff perform duties and digitisation without a sectional home. Finally, due to the complexity of oral history, creating metadata took a significant amount of time and effort. It was discovered that the metadata schema had some language problems because it was incompatible with other characters found in multilingual metadata descriptions.

6.2.5 Proposed metadata schema framework for oral history content

Several factors influenced metadata creation and access to oral history. The findings revealed several factors that could obstruct access to these collections. The following suggestions were made:

- The fact that the metadata schema was not compatible with some language characters clearly showed that the schema should be transformed and adjusted to allow flexibility in metadata creation.
- Translation is required for English and other languages.
- Backlogs in metadata creation and digitisation require more staff trained in technological skills.
- More training is needed to empower staff with metadata skills.
- AI is required to assist in metadata creation, especially in the content and language interpretation of oral history content.
- There is a need to revisit the legislation for archival institutions to accommodate and ensure metadata accountability for oral history content.
- Most importantly, the budget was a big challenge, as it affected all the recommendations to be implemented.

These recommendations strengthen the need for the proposed framework for access and discoverability of oral history content. The next section focuses on the conclusions drawn from the summary as discussed.

6.3 CONCLUSIONS FROM THE FINDINGS OF THE STUDY

The study should end with a statement about what the researcher discovered about his or her subject matter and where future research should be directed (Babbie 2013). Several conclusions can be drawn from an analysis of the content and interview transcripts. The primary research problem of this study was the need for an effective and efficient metadata schema to facilitate access to oral history content. Due to the state of the metadata at the NFVSA, the oral history content is technically hidden and inaccessible on the database. Because some of the created metadata remained hidden in the database, some oral history content was available but could not be located, and the oral history content items were not linked to each other. Some content may appear on the system, but cannot be downloaded or used. As a result, the oral history content at the NFVSA is largely inaccessible. Failure to strike a balance between the metadata schema and accessibility of oral history content would jeopardise the NFVSA's mandate as a national entity in South Africa, as the oral history content will be inaccessible and underutilised. This would make it difficult for communities and researchers to locate oral history content and African knowledge would remain unknown to South Africans, while communities would be unable to hold the government accountable for addressing past injustices. Researchers would also either not complete their research projects or experience unnecessary delays as a result of incomplete information due to oral history not being linked to each other. South Africa is working towards transformation and decolonisation, and citizens must understand their roots to have a meaningful impact. Without access to oral history, this would be hampered and/or impossible. This will also be difficult to achieve if there is a mismatch between the curation/collection process and the metadata schema used to describe oral history content. The incompatibility of the metadata schema for an effective description of oral history would affect not only the country's social and economic activities, but also its knowledge production. South Africa will make a smaller contribution to the knowledge economy, not only in the country, but also in Africa and globally. As a result, the primary objective of this study was to investigate the metadata schema for oral history content in order to propose a metadata framework for discoverability and access to oral history content. The

framework is presented and discussed as part of the recommendations section (see section 6.5 in this chapter).

6.3.1 Metadata schema used for oral history content at the NFVSA

Metadata schema is required for the standard description of any record. In general, the metadata schema must be adaptable to accommodate different incoming data types and their relationships in the repository. For example, the description should consider the language and script, who created the metadata and when, as well as the vocabularies. The consideration of data types strengthens the relationships between metadata records, regardless of their nature or the technical standard upon which they are based (Vos 2013). This means that when metadata are created, different types of metadata are considered. Taking into account the metadata process and elements used at the NFVSA, three types of metadata are being created: administrative, descriptive, and structural. As a result, one can conclude that the NFVSA values metadata quality and only requires the metadata schema that supports metadata creation. In simple terms, metadata categories are used to maintain metadata quality and metadata management (Phillips 2015). Therefore, all elements were considered to ensure relationships between metadata records (Park & Tosaka 2010; Mazé 2014), despite some not being used or information on the metadata records remaining incomplete.

Due to their activities, a comprehensive metadata infrastructure should account for them. In view of the various metadata entries deposited in national archives, metadata for the same infrastructure should be created at different times. As a result, the NFVSA metadata schema should be combined with other schemas and upgraded to meet additional requirements for the oral history content.

6.3.2 Metadata standards for oral history content

A useful metadata standard for oral history collections should provide a structured and consistent framework for describing these materials. This implies that care must be taken during the creation and curation of any oral history. In future, archivists should work towards ensuring that metadata are correctly processed so that oral history content can be accessed. Although metadata standards are used to support content access, they should also be used to ensure system interoperability and oral history discoverability (Matusiak, Werling, Donovan & Carlson 2021). However, to increase flexibility and detail in describing oral history

collections, it may be beneficial to use metadata standards and/or a customised metadata creation (Gill, Vos, Villa-Torres & Ramirez 2019). Flexibility would have helped to apply metadata standards in a reasonable manner during the migration of metadata records from Dublin Core to AtoM. Despite the fact that records were migrated from Dublin Core to AtoM, accurate and complete metadata are critical to migration and because Dublin Core was not fully compatible with the AtoM, data clean-up became a concern. ISAD(G), ISDIAH, and ISAAR were used and were embedded in AtoM, and provided guidelines on how to provide relevant data for metadata elements found in AtoM. This was also due to the staff's familiarity with these standards, despite the fact that they still required training.

6.3.3 Access to oral history content in the NFVSA

Oral history is unquestionably the foundation of African history in general, and South African history in particular. According to Zimu-Biyela (2022), preserving and promoting access to oral history records in South Africa aids in the development of community identity based on its storylines. As a result, oral history content must be accessible and discoverable to communities. Unfortunately, the findings indicate that, despite the need for access to oral history content, there are still metadata and technological barriers to accessing it. Users are still required to go to the NFVSA to access the materials. On the other hand, some oral content cannot be accessed due to a lack of metadata. As a result of the delay in digitising these oral history materials and the lack of metadata, there are no digital objects. The fact that the NFVSA continues to operate the old NAAIRS alongside the new NAAIRS is a challenge in itself.

On the other hand, the access points were not clearly linked in the finding aids. It was found that, in some cases, there were no inventories. Furthermore, the oral history content was restrictive, imposing certain constraints on the creation of links on the database. This means that even in cases where the metadata record permitted access, some of these contents were not easily accessible because no meaningful and useful access points to these recorded collections (both audio and audio-visual or a combination thereof) were created. As a result, despite the NFVSA's possession of these recordings, many of the oral history content may be inaccessible. Therefore, a framework is required to ensure access to oral history.

6.3.4 Technologies used for oral history content at the NFVSA

New technologies are enabling unprecedented global access to, and knowledge of, the world's recorded heritage (Frisch 2016). Simultaneously, these technologies are facilitating the generation, dissemination, and multiversioning of a wealth of potential by making metadata searchable via database (Crowe & Meagher 2015). A database like the NAAIRS is designed to make oral history content more accessible. Users would be able to search the database for information by the interviewer, interviewee, date, collection, and description. The old NAAIRS is still operational and operates alongside the new one because some metadata records are not available, and the new NAAIRS is not accessible to public users. Secondly, there was a backlog in the creation and digitisation of metadata. As a result, the use of AI could speed up the process for both.

The fact that there is system downtime and network challenges that limit metadata creation clearly demonstrates that the NFVSA should prioritise financial resources to maintain its systems and ensure full metadata schema implementation.

6.3.5 Proposed metadata schema framework for oral history content

Implementing the proposed metadata framework will ensure that the NFVSA improves the discoverability and retrievability of oral access content and related content as a way to address current challenges. The proposed framework will provide metadata access, allowing links between items within the same archival group as well as between archival materials. The framework will also ensure that the search functions are user-friendly and that users can access the oral history content without physically visiting the NFVSA site. Furthermore, the framework will ensure that the oral history content is not technically lost within the database and is easily found and accessible. The metadata framework will ensure that the time-consuming and labour-intensive process of creating metadata is offloaded and the backlog is significantly reduced. It will also increase the possibility of fast-tracking digitisation to ensure that the metadata match and are linked with the item in a format that users can access. Most importantly, despite its complexity and uniqueness, the oral history content will be effectively described, and the metadata schema used will be improved to better fit the purpose. This study investigated, developed, and recommended an embedded framework that can be used in the NFVSA to ensure access to oral history metadata. This framework will add best oral history practices that will assist the NFVSA in producing quality metadata and curatorial access in

South Africa to achieve its national mandate and ensure that African knowledge is sustained and South Africa remains a better competitor in knowledge production, both locally and internationally. Figure 6.1 in section 6.5 of this chapter shows the proposed framework.

6.4 RECOMMENDATIONS

Given the research problem, the metadata for oral history content and the metadata schema must be reviewed and improved to ensure effective access to oral history content at the NFVSA. This section explains the recommendations based on the objectives of this study.

6.4.1 Metadata schema of oral history content

The objective of this investigation was to describe the schema of metadata for oral history at the NFVSA. This section provides recommendations for the NFVSA metadata schema. It is proposed that, at the NFVSA, metadata records should be created and maintained in digital format. These can be saved in the cloud storage facility, allowing for unlimited access to the contents. This would then ensure that digital records could be accessed through a database, allowing users to browse through records remotely. However, the NFVSA should ensure that only authorised individuals have access to their cloud storage facilities and that ethical issues regarding access to oral history content are followed. The institution may think about incorporating ethical clearance procedures into the system. To improve access to oral history content, digital records should be retrieved by typing the reference number, subject, and keywords into the database.

6.4.2 Metadata standards for oral history content

There is no single metadata standard that can adequately describe all the documents emerging in various formats and aid in efficient information retrieval. There is an urgent need to develop a single unique metadata standard in order to provide the emerging global information infrastructure with coherent methods of organisation and access that transcend historical borders of nations, languages, and cultures (Ramesh, Vivekavardhan & Bharathi 2015). Since the results showed that there is a need for exposure of locally created metadata and metadata guidelines, locally added metadata elements will continue to be important in addressing the unique nature of oral history content. The problem arises when interoperability is not possible and collections are not sharable across the system and/or oral history collections. As a result, this study suggests that AtoM metadata elements be used with flexibility to allow metadata

created to align with current standards. Some of the elements that should be prioritised include "content," which allows archivists to create subject access and interoperability of dictate from one file to another, as well as between systems across provinces and South Africa in general. This means that systems must be monitored to ensure that mandatory elements are used and that metadata records are filled in.

As a result, the NAROS should be monitored and updated more frequently. This will ensure that the oral history content register and metadata creation at the NFVSA are captured, and that any missing data are identified in a timely manner. The NFVSA should implement measures such as human resources, training, and policies based on the recommendations in section 4.5 on the metadata framework. The adoption and/or review of metadata standards should be guided by the archival institution's policies.

According to the findings of this study, metadata interoperability remains a major challenge of interoperability across digital repositories and collections when there are no sharable mechanisms for locally defined extensions and variants.

6.4.3 Access to oral history content

Metadata can be stored in a database and linked to the original resource in some way, or it can be embedded in (or otherwise directly associated with) the original resource. Metadata for resource discovery and rights management could be included in a searchable database that provides access to oral history content. Metadata, on the other hand, should specify the technical formats used. The search function should include a clear indexing word, and the search function should be simple and accessible, regardless of location. The system should allow users to enrich and filter metadata while maintaining vocabulary control.

6.4.4 Technologies used for oral history content

This study proposes recommendations for the management of metadata records at the NFVSA. It suggests that metadata records should be maintained in digital format and accessed through a database search using reference numbers, subjects, and keywords. This search can use Boolean logic and filters for efficient retrieval.

A key question raised is the level of detail required for the description of archival records. Should the descriptions be at unit level or go deeper into specific items? If more detailed

records are desired, innovative solutions would be necessary to streamline the description process while keeping it concise. Additionally, the creation of indexing thesauri and request compilations is suggested, although the feasibility of this undertaking depends on the potential benefits it brings.

The study also acknowledges the changing nature of historical sources as digital archives become increasingly important. These born-digital archives pose unique challenges for digital preservation practitioners, archivists, and researchers. Issues such as data volume, hardware and software obsolescence, standards, and contextual changes complicate the preservation of born-digital materials. The diversity and historical fluidity of digital materials create challenges for data recovery and analysis, which require specialised knowledge and tools for forensic analysis and interpretation . Shibambu (2023) also emphasised the importance of digital storage to assist obsolesce structures in archives.

Ethical considerations, the preservation of online communities, and the recovery of archived websites also pose challenges for digital archivists. Future trends suggest that technological advancements may lead to partially automated archival description, including repurposing existing metadata. Institutions lacking policies, infrastructure, and trained personnel require immediate attention.

The study recommends training programmes for staff that focus on developing metadata skills and expertise in AI and oral history content. The use of AI infrastructure and staff proficiency are crucial to effective access to distributed digital repositories and oral history collections.

In conclusion, the study recommends improving metadata management practices through digital formats and sophisticated search methods. It emphasises the importance of training staff in metadata skills and technological tools, including AI, to address the evolving landscape of digital archiving.

6.5 PROPOSED METADATA SCHEMA FRAMEWORK FOR ORAL HISTORY CONTENT

The purpose of the study was to explore a metadata schema for oral history content and to propose a metadata schema framework for accessibility and retrievability. There is no widely accepted guide for metadata on oral history content. Where guidance does exist, it has been influenced by metadata from non-book materials, particularly practices developed for records

management and other archival materials. Based on the conceptual framework, the objectives of this study and the data collected and discussed, the researcher proposes the metadata framework depicted in Figure 6.1.

The framework takes into account the country's lack of legislation specifically geared towards oral history collections, but acknowledges the best practices outlined. The researcher also considered the complex nature of oral history content and the lack of standardisation of archival practices in the country as a repository of oral history content. The framework is a detailed description of the processes, resources, and other factors that influence the success of metadata and access to oral history content.

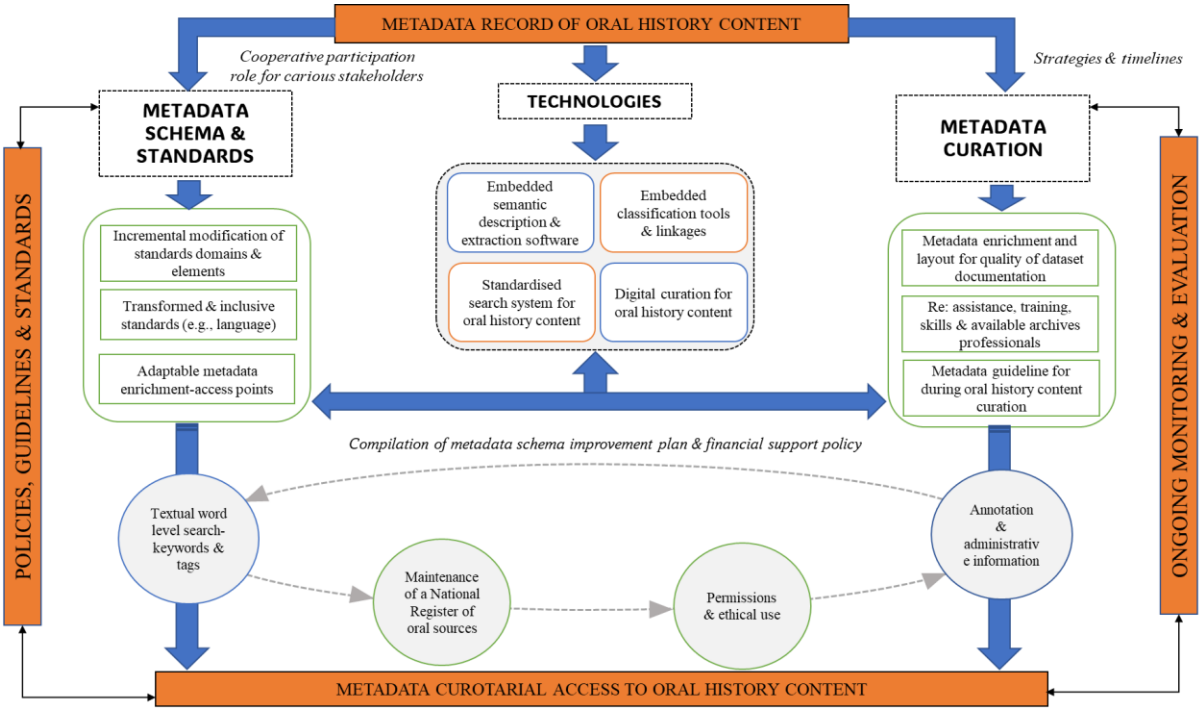


Figure 6.1: Framework for metadata and access to oral history in archival institutions (Researcher' own work)

To begin with, the framework displays a collection of schema improvement plans based on metadata.

(a) Policy and guidelines document availability

Before implementing any plan, it is critical to determine whether the project aligns with existing policies.

review/develop and/or implement policies related to the intended plan. Policies pertinent to this framework would include:

(b) Technologies for metadata schema, standards and metadata curation

Preservation and accessibility of oral history recordings remain a critical goal in archival institutions. The NFVSA helps to ensure that these important historical resources are available for future generations to study and learn from by providing a centralised repository for these materials and promoting best practices in their collection, preservation, and access. Given the nature of oral history content, it is mostly time-consuming while sometimes complex to create the metadata. Technologies will improve the quality of metadata and reduce the workload associated with metadata creation for oral history content. This suggests an improved technological approach to embedded systems to ensure metadata quality to ensure access to oral history content.

An embedded semantic description and extraction software tool is needed. Looking at the format of oral history content, it may take time for archivists to listen to audio and record metadata. Besides the time spent on one item, there is also the mental capacity level required for each case. Hence, metadata creation is always on a backlog. It is thus significant to embed a software that will do text analysis and interpretation, and extract keywords and/or text word for access. Oral history access could also be searchable using keywords that link related oral history content. This would enable users to search for and discover content without knowing the title or creator/author. The improvement of text word-level search for oral history content in archival institutions increases not only accessibility, but also awareness and visibility of the essential function that archival institutions play in holding the country accountable.

It will also be more beneficial to have an automated metadata system. If this happens, the metadata management will be simplified. Duplicates will be detected easily, while incomplete and/or missing elements will be identified. This will also allow linkages to avoid items that are identified as in the system but cannot be located, as such items are technically lost. Links will help link metadata for oral history content to the content, but also to others related to the main metadata. It is worth lauding the NARSSA that it has taken considerable strides to embed a standardised search system, even though it is not fully functional. Furthermore, there is a continuous need to maintain the metadata schema and standards used at NFVSA.

(c) Metadata schema and standards

This element of the framework covers refers to metadata localisation and customisation. By localising and customising involves modification of standards that would be inclusive and compatible with various languages as well as adaptable access points. Metadata should be incremental or adaptable to suit the collection at hand. It should provide the flexibility required by the community to meet their needs in terms of discovery and access, and long-term sustainability, including the selection of locally relevant terminology and vocabularies, as well as the creation of custom elements. Balancing localisation/customisation can pose challenges to interoperability and sustainability, but these can be overcome if the community articulates their ultimate goals for the project clearly. Throughout the literature, there was a strong community desire to be able to contribute metadata in the form of tags or keywords, descriptions, or stories.

Depending on the institution and the needs of the collection, the NFVSA can also use other metadata standards or local metadata schemas to describe oral history collections. For example, the United States Holocaust Memorial Museum describes its oral history collections using a customised metadata schema that includes elements such as the name of the interviewer, the name of the interviewee, and a detailed description of the interview content. Overall, metadata are an important aspect of the NFVSA's and other institutions' oral history collections, as they help ensure that these materials are properly documented and organised, and that they are accessible to researchers, scholars, and the general public.

It is critical to ensure that certain characters are captured accurately when using the metadata schema system to capture metadata for oral history collections. Special characters, diacritical marks, and non-English characters, which are commonly found in oral history materials, are examples of these characters. Archivists and researchers should use encoding systems such as Unicode, which supports a wide range of characters and diacritical marks from different languages and scripts, to accurately capture these characters. They should also ensure that their metadata capturing systems are set up to handle a variety of character sets and that their data entry procedures include instructions for capturing and encoding these characters. Furthermore, archivists and researchers should be aware of any local or national character encoding standards and ensure that their metadata adhere to these standards. This will aid in ensuring that their metadata are compatible with other systems and can be shared and reused by other researchers

and institutions. Hence, metadata curation remains another important factor in ensuring that oral history content will be accessible.

(d) Metadata curation

Metadata curation involves metadata enrichment, skills and metadata guideline for oral history content. As already indicated that the metadata schema should be adaptable to metadata enrichment. This suggests that the metadata should be adaptable to adding more quality access points to accommodate the needs of the users. The layout should be accommodative of the new developments in metadata creation without changing the original information. Metadata curation requires knowledge and skill training. The backlog of metadata has become a concern and one of the reasons cited was the limited number of staff to handle the workload and a need for training. The NFVSA may provide training and support to individuals and organisations interested in conducting oral history interviews in addition to collecting and preserving oral history recordings. This can include recommendations for interview techniques, equipment, and best practices to ensure the quality and accuracy of recordings. On the other hand, training may be provided on metadata schema to enhance its use and ensure quality metadata.

Lastly, training should cover not only metadata content and structure, but also the technological skills required to achieve metadata quality. It was found that some of the metadata backlog has emerged out of migration and slow-paced digitisation, which was a result of minimal staff. This called for increased staff. For example, only two staff members were assigned the task of collecting oral history. As such, the backlog begins with the delay to collect, and then the delay due to staff overload.

The increased number of staff will also help the increase of staff for digitisation. There is a need to establish a digitisation unit to offload other sections with digitisation workloads. This will also impact the trainings to be provided. A need for staff capacity and development will have a very high bearing on policy and policy makers because they require a revision of organisational structure and job descriptions to that effect. Therefore, policy review and development form the basis of this framework.

(e) Cooperative participation for various roles

The metadata are created only once and then reused numerous times. As such, it should be carefully created to be inclusive of any data that would enhance discoverability and accessibility of oral history content. Cooperative participation becomes a crucial towards the successful metadata schema. This cooperative participation invoves collaborations that will benefit the metadata schema and metadata creation process. The collaboration of both internal and external stakeholders is crucial for the metadata access of oral history content.

Internal collaborations

Internally, metadata creation involves a separate unit at the NFVSA. For example, curators of oral history content, information management systems, and metadata specialists should work together. While knowledge on metadata schema is important, it can be enhanced by collaboration with information management system specialists. The problem around missing data and/or incomplete metadata records would be the result of different stakeholders working in silos. For example, an oral history collection might omit certain information required by the metadata archivist to create metadata.

Collaboration between provinces is also crucial. Metadata that are created at that level and deposited into the national archive should not be found as duplicates or inconsistencies on record for the same items. Metadata can also improve efficiency in terms of the timing of metadata capturing. There is also a need for coordinated oral history collection projects, thus metadata capturing. On the other hand, it was found that the implementation of the provincial mandate to deposit their metadata was not fully functional. Only three of nine provinces, at the time of data collection, were actively involved in metadata creation and carried out their mandate for metadata creation.

External collaborations

External collaboration with various stakeholders will help to learn best practices that can be adopted for metadata access for oral histories. For example, language experts might be plugged into a multilingual and language metadata system. There is a need for collaborative metadata management strategies that can provide efficient and cost-effective methods of acquiring and maintaining certain types of metadata, for sharing, re-use and disseminating metadata capturing throughout the metadata cycle. There are also archival institutions that are not subdivisions of

the national archives. Collaborations will increase the dissemination and access to oral history content.

These collaborations are in line with system interoperability and metadata management to ensure discoverability and access to oral history content. This may have major implications on the existing plan hence the need for improvement plan.

(d) Metadata schema improvement plan

The improvement plan incorporates permission and ethical use of oral history content, maintenance of the metadata system such as NAROS to ensure accessibility and retrievability of oral history content.

Some oral history content requires requesting permission for use. The request for permission should be embedded as part of the search strategy. Users must be aware of the nature and conditions of access for each collection before the decision to borrow it. This suggests that all types of metadata (administrative, descriptive, and structural) should be fully described and clearly linked to permissions and use of the oral history content. The NFVSA collects, preserves, and makes audio-visual materials, including oral history collections, available to the public. The metadata used to describe these collections vary from one institution to the other but typically include information such as the date and location of the interview, the names of the interviewer and interviewee, and the subject matter covered in the interview. This involves registration of these content in the archival register, NAROS.

It should be emphasised that the descriptions of the archival collections in the NAAIRS contain information on the archival holdings of the National Archives and Provincial Archives Services and incorporate those of the NFVSA (NARSSA). NARSSA is responsible for the maintenance of a NAROS, at which stories are registered. The purpose of this database management system is to give interested parties access to research findings on topics that could be of interest to them. Navigating through data, searchers find oral history collections. There should be a proper plan on how to allow navigating while ethical permissions are accounted for. The need for ongoing monitoring and evaluation becomes a requirement for metadata schema framework implementation.

(f) Ongoing monitoring and evaluation

Monitoring and evaluation should be an ongoing process, from the beginning to completion. This will ensure the efficiency and effective implementation of the metadata framework. Reviewing and tracking the progress and activities will improve ways in which metadata schema can be used to enhance discoverability and accessibility of oral history content. This will include policy review and development. The following policies might be considered:

- **Policy on transformation and colonisation:** Transformation and colonisation are central to South Africa's core business.
- **Metadata management policy:** Principles for oral history content metadata
- **Financial support policy:** Funding is essential for any improvements to be implemented.

As a result, it is vital that the policy should clarify the process of transforming the metadata schema and standards, in addition to the archival institution. Guidelines for improving existing practices in archival institutions, particularly oral history, should be developed in conjunction with the record management policy. Since oral histories are kept in archives, they are treated the same way as any other record or archival material. The transformation policy will not only allow customising the metadata schema to best suit metadata for oral history content, but will also decolonise the archival access and description of oral history content. Knight (2022) makes a recommendation to "decolonise access and classification" and begins by examining the difficulties involved in addressing this recommendation while working within a colonial institution representing a colonial worldview. Transforming and/or decolonising the processes in the NFVSA may encounter challenges that may have an impact on metadata creation. One cannot ignore the general characteristics of Western and indigenous worldviews and the impact thereof on metadata and metadata schema for oral history content. Hence, the transformation policy should be planned and developed to enable archival institutions to also decolonise archival access and metadata creation.

On the other hand, there is a need for management of metadata, especially with the migration of the metadata to the new metadata schema. Migration of content may be easy or necessary to transfer data from one carrier to another for access or preservation, but doing so may result in the loss of contextual meaning and important information." (Edmondson 2016). Migration and/or change can also result in a change in content. For example, loss of image or sound

quality may change the content of oral history content. While enhancing the quality of metadata, provenance might be lost.

Metadata management is also important in the case where archival institutions should be able to create metadata within the context of South Africa. For example, the multilingual nature of the oral history content should be managed, as well as the different formats. Oral history content is collected in different languages, some of which may not be compatible with the metadata schema and standards used. This poses challenges, as the metadata may not be a representation of the collections, since some characters may not be captured or adapted, violating the principle of origin and provenance. Metadata management policy is a necessity to address factors that might arise during the process. These adaptations and management require financial capacity and budget reallocation to improve and/or customise the metadata schema; hence, the importance of financial policy. The policy should also specify the roles and responsibilities of various stakeholders. Although roles may be clearly defined, to be effective, stakeholders need to collaborate.

To summarise, when using the metadata system to capture metadata for oral history collections, it is critical to ensure that specific characters are correctly captured. This is possible with encoding systems, careful data entry procedures, and adherence to local and national character encoding standards. Metadata schema such as AtoM is a metadata encoding standard for archival materials, such as oral history collections. AtoM provides a set of standardised elements, such as title, creator, date, subject, and description that can be used to describe the content and context of archival materials. The metadata schema is intended to make archived materials more accessible and to help in their long-term preservation. In the context of AtoM, a translator is software or a tool that can be used to convert metadata from one format to another. Based on a set of rules or mappings between the two formats, the translator would read the original metadata in its original format and then convert it to the AtoM format. The resulting AtoM metadata could then be used to make oral history collections more discoverable and accessible.

6.6 IMPLICATIONS OF THEORY AND SCHOLARSHIP

Implications in research imply that researchers can forecast future usefulness or that a reader can relate to the usefulness of the research in some way in their specific context (Garcia 2018). Garcia (2018) further state that research implications help to inform, teach, and influence people's lives for the better. In some ways, implications refer to who, what groups, or what institutions, including NFVSA, should benefit from this study. This study will contribute to improved access to oral history in South Africa.

The purpose of this study was to explore a metadata schema for oral history content in South Africa in order to propose a framework for ensuring its access and retrievability. The findings of this study outlined important aspects of metadata in South African archival institutions. Most of the issues to be considered are related to metadata and improving access to oral history content, as well as being central to the effectiveness of discoverability and access to history content at NFVSA.

The study also made additional recommendations in response to the identified challenges. The recommendations of this study may result in an improved metadata schema and record appropriate for the nature of oral history content in South Africa. This is due to the recommendations' intent to address the challenge of creating metadata for oral history content, the backlog in metadata creation and digitisation, the lack of a user-friendly search page, the inability to translate, and items on databases with correct metadata but inaccessible due to incompatible technologies.

Addressing these challenges would require embedded AI, increased staff, and staff capacity. This will also help automate the description and classification of automated sound records. Metadata schema and standards should be integrated in such a way that they provide effective storage capacity and links, allowing users' requests to be addressed more effectively. Metadata technologies would also help users retrieve archival records more quickly, as well as scan and classify content for the purposes of linking content internally at the NFVSA and in South Africa in general. This would manage preservation metadata and act as a backup in the event of a system crash, resulting in the loss of metadata and other records.

The system should include ethical issues that must be followed to facilitate faster access to oral content and avoid potentially illegal, restricted, or lengthy access processes. This study also

recommends an embedded framework to help the NFVSA implement an integrated approach to metadata creation and oral history access.

The framework focuses on NFVSA metadata applications and technologies that will improve access to oral history content. In general, this study contributes to theoretical and practical application in the field of oral history and archival studies, and in particular to metadata and curation access to oral history. Based on the findings, recommendations, and framework introduced or proposed, the study will help the NFVSA establish and review its own policies, procedures, and best practises. It also directs the systematic implementation of best practices for metadata and information management systems.

All activities that must be performed to place metadata in the right space for the right user are included in the description and arrangement of oral history content. According to the study findings, the oral history content in the NFVSA is appropriately labelled and packaged. This implies that the metadata of oral history content allows for the sharing and facilitation of access to these contents. The database should include metadata to help searchers find relevant content in the archives. With finding aids, one should be able to discover, evaluate, and access whether the information retrieved meets their information needs. The study also revealed that the NFVSA has finding aids with detailed lists of content to allow faster access. However, the tool was found to be difficult to use. The fact that there were two databases running (one old and one new) as a result of the migration and digitisation backlog, is also a source of concern. The NFVSA is understaffed, which contributes to the backlog, and digitisation capacity adds to this difficulty.

The implication for remote access is that the NFVSA will require the NAAIRS search page to be customised according to user needs and that the migration is complete in order to allow access to the new database, which appears to be more user-friendly. The need to extend remote access is also influenced by the fact that the NFVSA's operating hours were limited; therefore, digital adoption was required. According to this study, the Archives are always closed to searchers on Saturdays and public holidays.

According to the study findings, the indexes were not available in NFVSA, which presented difficulties because users must be directed to the appropriate source and have quick access to it. Filters are used in the finding aids, implying that users should know exactly what they are looking for. Another issue that arose was the difficulty of linking content. As a result, the study

suggests that the metadata schema be used in conjunction to ensure subject organisation and a comprehensive index, allowing for discoverability, retrievability, and access to oral history content.

This research will make an original contribution to knowledge, including recommendations for the operationalisation of relevant metadata elements in metadata activities in South Africa. In addition to exploring the challenges faced by archivists in carrying out the activity, a deeper understanding can be initiated by knowing the effects that contributed metadata schema has on metadata access of oral history content. In doing so, it is hoped to create a new dynamic environment to enhance metadata access to oral history content in South African archival institutions.

Oral histories are notoriously difficult for archivists to work with. In terms of description, copyright, and overall management, they are complex and distinct from other types of archival records (Chernevych 2016). The audio-visual media recordings are interactive and non-standardised, and they frequently lack key metadata or vital associated records. These are only a few of the difficulties that archivists face when processing and caring for oral histories. Navigating these difficulties necessitates expert advice, best practices, and templates. This research could provide much needed assistance.

Simply put, providing access to oral history content in archival institutions requires a commitment to continuous improvement. In anticipation of growth in oral history content, metadata access and issues are likely to impact access to oral history content, as well as technologies and work environment.

More research should be done on the issues associated with metadata, metadata schema, and metadata creation, and, in general, providing access to oral history content. Hence, the next section focuses on recommendations for further research.

6.7 RECOMMENDATIONS FOR FURTHER RESEARCH

This study explored the metadata schema for oral history content in South Africa. While considerable attempts have been made, there are several key areas where more studies are needed if archival institutions want to maintain access to oral history content in South Africa. Among these are the following.

6.7.1 Information search patterns for the archival institutions

User search patterns, whether physical or virtual users, may assist archivists in noting and observing commonly search files that would require to be enriched metadata for future searching exercises by other users and/or oral history content seekers in the archival environment. All users should be aware of the policies, procedures, and tools for managing records and they should be capable of applying them consistently to all records. The full cooperation of the users is necessary to file documents into the filing system and to protect records against loss and damage. Without this, there will be no records to manage. Hence, such a study is recommended.

6.7.2 Collaborative approach to oral history collections in South Africa

This study revealed that while numerous oral history projects are under way in the country, little is known about the oral history collections being collected by various stakeholders, such as academic institutions and OHASA. OHASA, as NARSSA's brainchild, was supposed to report and deposit its content to NARSSA, which does not appear to be done. OHASA has held annual conferences in various provinces of the country, with research having been conducted and published. Little is known about oral projects undertaken for the purpose of accessibility. OHASA is a product of the Department of Arts and Culture and the post-1994 political transformation. The seed that gave birth to OHASA was planted in 1999, when Cabinet mandated the then Department of Arts, Culture, Science and Technology (DACST) to conceptualise and spearhead the National Oral History Programme (NOHP) for South Africa (OHASA 2023). A tracer study would be useful not only for OHASA, but also for other stakeholders involved in oral history projects.

Collaboration can make it easier for users to access oral history content services. Collaboration can also "make core services more convenient for users, reduce workload for individual institutions, and increase standardisation of policies and digital platforms" (Masenya & Ngulube 2021). Public libraries, for example, are well positioned to provide services to community members that are easily accessible and within reach. In South Africa, archival institutions are mainly concentrated in the mother cities of each province. Research into collaboration between archival institutions and public libraries could be conducted.

6.7.3 Oral history metadata schema in academic institutions

Metadata improve access to oral histories but raise new concerns for those who collect them. Oral history content, as a common form of documenting the past, is found in the collections of many academic libraries, according to Grimsley and Wynne (2009). According to the NARSSA database, academic institutions such as the University of Pretoria and the University of Cape Town were members prior to the adoption of the new metadata schema. A study could be conducted using examples of metadata records established at these institutions to outline practical strategies that academic libraries can use as a guide in managing access to oral histories, and it could present some solutions to these problems in an effort to increase standardisation among academic libraries and archives. A study could be carried out to determine the prevalence and scope of oral history collections in academic institutions as well as why some institutions collect oral history content, while others do not. This would involve conducting a tracer study of oral history collection in South Africa.

6.7.4 Legal framework for oral history content

According to this study, there is no standard guideline and/or procedure in the country to coordinate access to, and accountability for, oral history collection. There appears to be policies and guidelines in place regarding archival description and access to archival collections in general. A study aimed at developing policies and procedures for collecting, acquiring, managing, and disseminating oral histories is highly recommended. Some research questions could include, “How many institutions have policies and/or procedures in place for the collection, processing, and/or use of oral history content?” and “How far do these policies go in terms of addressing metadata for oral history content, as well as accountability for making it accessible and preserve it?”

6.7.5 Digital transformation in archival services

The primary function of any archival institution is accessibility. To accomplish this, embracing the use of AI in archives will benefit not only metadata creation and records, but also the preservation of these collections. A study to investigate the application of AI in the creation of metadata for oral history content is highly recommended. This study found that the lag in metadata is due to the amount of time and mental capacity required to create metadata for this oral history content. For instance, the amount of time and concentration required to play and

listen to oral stories in order to identify elements required for metadata creation. The study may seek to identify and investigate the available AI tools to help with metadata. This study recommended the use of AI for metadata creation at the NFVSA, but additional research would be recommended for the use of AI for metadata at other South African institutions that collect oral history content.

Archival institutions play an important role in providing access to oral history content and other archival materials in any country; therefore, they should have the resources to help them successfully fulfil their mandate of providing citizens, researchers, and other institutions/organisations with access to oral history, and to provide efficient support and services. It is critical to conduct an analysis of CC techniques to identify significant changes in metadata creation and records in various formats of oral history content. Furthermore, an assessment and/or implementation of AI in the preservation and access of oral history content in digital formats is recommended.

6.7.6 Transforming metadata schema system in archival institutions

Archival institutions should develop strategies to transform and incorporate the metadata produced into archival and preservation metadata schemas.

To be truly effective, metadata schemas are designed with retrievability and accessibility sustainability in mind. While archival staff appear to be more satisfied with the system, there is still room for improvement and closing gaps where necessary. The metadata schema has also taken a step forward in ensuring its translation into various languages around the world. However, not all languages are included, and there are no clear guidelines on the translator's requirements. On the other hand, little is known about the project's successes and failures. It is thus necessary to conduct an analysis of the AtoM as the metadata schema elements, as well as an analysis of the translation progress, to identify success, and, where possible, to develop a framework to improve the project's effectiveness. The role of communities in the creation and development of such tools should also be investigated.

6.7.7 Digitisation in archival institutions

Archivists use a variety of techniques and methods to protect and maximise the use of archival materials. One of these techniques is digitisation. The digitisation of archival collections broadens access to rare historical materials and creates a critical mass of primary sources that

can be used for research, education, and personal purposes (Matusiak et al. 2021). It allows maximum use and protection of archival materials. However, aside from the benefits of digitisation, it is a technical process. Thus, experts should plan and carry out digitisation, taking into account factors such as archival material and users. Moving from analogue to digital requires effective restructuring of the relationship between the institution and archival materials. Digitisation can happen for a number of reasons. Thus, it is necessary to investigate the nature of digitisation and conduct a more in-depth analysis of user experiences and oral history content in relation to archival institutions.

6.7.8 Database searches and access to archival materials

The way users remember, search, and use archival materials should guide the design of search tools. This requires interoperability, linkages, and metadata provenance for oral history content (Taniguchi 2023). Further research could be conducted to investigate commonly used access methods and their application to oral history content. This will influence whether there are differences or a standardised approach to providing access to oral history by various organisations/institutions. Even systems would follow various metadata management policies. Given this, research is required on the preparation, transmission, receipt, and information packages for archived content and metadata within a system and across multiple system environments. Such research would shed light on the effects of inter-repository transfer of metadata and content as it moves into and out of a digital archiving system.

6.7.9 Decolonisation and transformation in archival institutions

The metadata schema and standards were developed under the principles of western philosophies, that the language is not Africa orientated, and that the lack of standardisation is the result of issues such as colonial languages being adopted in the system, according to the literature. A study should be carried out to investigate the feasibility of decolonising and transforming schema/standards and archival institutions. The impact of colonisation and the progress made in advocating for transformation are topics for future research. On the other hand, the oral history content could be evaluated in order to determine the extent and direction of the collection's coverage, as well as how to improve the way forward.

6.8 FINAL CONCLUSION

It was clear from the study that the NFVSA has made some strides in the metadata for oral history content. However, there are still some gaps in the metadata access of oral history that should be filled. The fact that the database and/or website search function is not user-friendly, non-linkage of some items within the system, technical loss or displaced metadata records and oral history content items within the system, external inaccessible metadata, and oral history collections compelling users to physically be on site, record language and translation of other oral history items, backlog in digitisation, system downtime, and network connectivity are all examples of metadata gaps. The study made several recommendations for improving the metadata schema for oral history content at the NFVSA.

This study proposed a metadata framework to improve metadata access for oral history content in order to ensure effective metadata management of oral history content at the NFVSA. This could also result in faster digitisation of oral history content for discoverability and access purposes. To reduce the time required to complete each metadata for each item, the system could also borrow some AI technologies for embedded translations and audio text analysis. A system like this would also provide effective backup of metadata records and oral history items, while elements are extensible to ensure effective migration with less data clean-up and/or displaced records. For better access to archival materials, the NFVSA should have a proper information retrieval system as a national entity.

Researchers should not have to struggle to retrieve and access digital records, so the NAAIRS database will be improved to provide a faster and easier-to-use search function. It is critical for the NFVSA to preserve and create oral history content so that South African citizens can use and learn from it. On the other hand, as a research hub for accessing and contributing to research in the oral history space, the NFVSA would make a significant contribution to the country's knowledge production. Oral history content that is accurate and manageable would be discoverable and duplicates would be avoided. Instead of duplicates occurring within the NFVSA and across the country, the limited resources would be used to conduct more oral history projects. Policies and guidelines would also be reviewed and improved to ensure that metadata and related issues are properly accounted for, as well as their implementation strategies. This study helped to benchmark the current metadata schema regarding oral history content in South Africa against the global use and process on the topic at hand. It has also

become evident that inventories and metadata records can impact oral history in the future and need further study towards a more simplified implementation of metadata schema and standards. Metadata schema and access to oral history need more attention in South Africa, and awareness towards the metadata framework should be created among creators of metadata for oral history content.

Simply put, the objectives of the study were met and a framework for metadata and access to oral history content was proposed based on a review of the literature and findings derived from qualitative data collected and analysed. This research contributes to the metadata schema and metadata access gaps in South African archival institutions.

REFERENCES

Adom, D, Joe, A-A & Hussein, EK. 2018. Theoretical and conceptual framework: Mandatory ingredients of quality research. *International Journal of Scientific Research* 7(1): 438-441.

Agwor, TC & Osho, AA. 2017. Empirical reflection on business ethics and entrepreneurial success in SMEs in Rivers State, Nigeria. *International Journal of Advance Engineering and Management Research* 2(6): 1956-1965.

Ahmed, S. 2018. Seeking information from the lips of people: Oral history in the archives of Qatar and the Gulf region. *Archives of Science* 18: 225-240. https://doi.org/10.1007/s10502-018-9293-8

Akaranga, SI & Makau, BK. 2016. Ethical considerations and their applications to research: A case of the University of Nairobi. *Journal of Educational Policy and Entrepreneurial Research* 3(12): 1-9.

Albarelli, G & Starecheski, A. 2013. The telling lives oral history curriculum guide. *New York: The Columbia University Center for Oral History*.

Alemu, G & Stevens, B. 2015. *An emergent theory of digital library metadata: Enrich, then filter*. Amsterdam: Chando Publishing.

Alemu, G, Stevens, B, Ross, P. & Chandler, J. 2014. Toward an emerging principle of linking socially constructed metadata. *Journal of Library Metadata* 14(2): 103-129. http://doi.org/10.1080/19386389.2014.914775

Alemu, G. 2016. A theory of metadata enriching and filtering. *Libri,* 66(4): 251-262. http://dx.doi.org/10.1515/libri-2016-0109

Alfier, A & Feliciati, P. 2017. Online Archives for Users: Premises for a Quality Reference Model". *JLIS.It* 8 (1):22-38. https://doi.org/10.4403/jlis.it-12269.

Allen, M & Cervo, D. 2015. Multi-domain master data management: Advanced MDM and data governance in practice. Amsterdam: Morgan Kaufmann.

Alma'aitah, WZ, Talib, AZ & Osman, MA. 2020. Opportunities and challenges in enhancing access to metadata of cultural heritage collections: a survey. *Artificial Intelligence* 53, 3621–3646. https://doi.org/10.1007/s10462-019-09773-w

Al-Saadi, H. 2014. Demystifying ontology and epistemology in research methods. *Research gate 1*(1):1-10.

Arifin, SRM. 2018. Ethical considerations in qualitative study. *International Journal of Care Scholars* 1(2): 30-33. https://doi.org/10.31436/ijcs.v1i2.82

Arnold, L & van der Walt, T. 2018. Collecting and Memory: A study of travel archives. *Journeys* 19(1): 52-66. https://doi.org/10.3167/jys.2018.190103

Arnold, LJ. 2015. *Refining metadata for genre archives: A case of travel archives.* University of South Africa, Pretoria. Available at: http://hdl.handle.net/10500/20119 (Accessed 27 July 2023)

Ary, D, Jacobs, L & Razavieh, A. 2009. *Introduction for research in education*. 8th ed. Belmont, CA: Wadsworth Cengage Learning.

Assmann, I. 2009. International Association of Sound and Audio-Visual Archives (IASA). In *Encyclopaedia of Library and Information Sciences.* 3rd ed. (pp. 2856-2860). CRC Press.

Aurini, JD, Heath, M & Howells, S. 2016. *The how of qualitative research: Strategies for high quality execution.* London: Sage.

Avery, JM. 2010. The democratisation of metadata: Collective tagging, folksonomies, and Web 2.0. *Library Student Journal* 5(8). Available at: https://www.librarystudentjournal.org/index-php/lsj/article/view/135/268.html (Accessed 16 September 2021)

Babbie, ER. 2013. *The practice of social research*. 13th ed. Belmont, California: Wadsworth Cengage Learning.

Baca, M. (ed.). 2016. *Introduction to metadata*. 3rd ed. Los Angeles: Getty Publications.

Baptista, A, Frick, L, Holley, K, Remmik, M, Tesch, J & Âkerlind, G. 2015. The doctorate as an original contribution to knowledge: Considering relationships between originality, creativity, and innovation. *Frontline Learning Research* 3(3): 55-67.

Batley, S. 2014. *Classification in theory and practice*. 2nd ed. Amsterdam: Chandos Publishing. https://doi.org/10.1016/C2014-0-00625-9.

Bell, E, Bryman, A & Harley, B. 2022. *Business research methods*. 6th ed. Oxford: Oxford University Press.

Bender, S, Hausstein, B, Hirsch, C & Bundesbank, D. 2018. An introduction to INEXDA's metadata schema. Ninth IFC Conference on "Are post-crisis statistical initiatives completed?" Basel, 30-31 August 2018

Berry, L. 2018. Review of "Curating oral histories: From interview to archive" by Nancy MacKay. *Oral history review,* 45(1): 151-153. https://doi.org/10.1093/ohr/ohy007

Bertram, C & Christiansen, I. 2014. *Understanding research: An introduction to reading research*. Pretoria: Van Schaik.

Bertram, C & Christiansen, I. 2017. *Understanding research: An introduction to reading research*. Pretoria: Van Schaik.

Bettivia, RS. 2016. The power of imaginary users: Designated communities in the OAIS reference model. *Proceedings of the Association for Information Science Technology* 53: 1-9. https://doi.org/10.1002/pra2.2016.14505301038

Beyene, WM & Godwin, T. 2018. Accessible search and the role of metadata. *Library Hi Tech* 36(1): 2-17. https://doi.org/10.1108/LHT-08-2017-0170

Bhebhe, S & Ngoepe, M. 2021. A forgotten past is the past that is yet to be: evaluation of oral history programme of the Oral History Association of South Africa. *Journal of the Eastern and Southern Africa Regional Branch of the International Council on Archives* 40: 60-78. https://dx.doi.org/10.4314/esarjo.v40i1.5

Bhebhe, S & Ngoepe, M. 2022. Building counter-archives: Oral history programmes of the Sinomlando Centre and Memory Work in Africa and the South African History Archive. *Information Development* 38(2): 257-267. https://doi.org/10.1177/0266666921999754

Bhebhe, S. 2015. Description of the oral history programme at the National Archives of Zimbabwe. *Oral history Australia Journal* 37: 49-55. https://doi.org/10.3316/informit.519672012616795

Bibi, H, Khan, S & Shabir, M. 2022. A Critique of research paradigms and their implications for qualitative, quantitative and mixed research methods. *Webology 19*(2): 7321-7335.

Bilushchak, T & Fedushko, S. 2016. Information and communications technologies in the archives and museum activities. *Proceedings of the 5th International Academic Conference ICS – Information, Communication, Society*, 19-21 May 2016, Lviv – Slavske, Ukraine.

Blaikie, N & Priest, J. 2019. *Designing social research: The logic of anticipation*. 3rd ed. Cambridge, MA: Policy Press.

Bless, C, Higson-Smith, C & Sithole, SL. 2013. *Fundamentals of social research methods: An African perspective.* 5th ed. Cape Town: Juta.

Bonache, J & Festing, M. 2020. Research paradigms in international human resource management: An epistemological systematisation of the field. *German Journal of Human Resource Management* 34(2): 99-123.

Borek, A, Parlikad, AK, Webb, J & Woodall, P. 2014. *Total information risk management: maximizing the value of data and information assets*. Waltham: Morgan Kaufmann.

Bos, J. 2020. Confidentiality. In *Research ethics for students in the social sciences.* Springer, Cham. https://doi.org/10.1007/978-3-030-48415-6_7

Bosum, A & Dunne, A. 2017. Implementing the Brian Deer Classification System for the Aanischaaukamikw Cree Cultural Institute. *Collection Management* 42(3-4): 280-293. https://doi.org/10.1080/01462679.2017.1340858

Bountouri, L. 2017. *Archives in the digital age: Standards, policies, and tools*. United Kingdom: Chandos.

Bourgeois, DT, Smith, JL, Wang, S & Mortati, J. 2019. Information systems for business and beyond. Saylor Academy.

Boyd, D. 2011. Achieving the promise of oral history in a digital age. In *Oxford handbook of oral history* edited by Donlad A Ritchie, pp.285-302. Oxford: Oxford University Press.

Boyd, D. 2013. OHMS: Enhancing access to oral history for free. *Oral history review* 40(1): 95-106. https://doi.org/10.1093/ohr/oht031

Boyd, DA & Larson, MA. (eds.). 2014. *Oral history and digital humanities* (pp. 141-162). New York: Palgrave Macmillan.

Bradley, K & Puri, A. 2016. Creating an oral history archive: Digital opportunities and ethical issues. *Australian Historical Studies* 47(1): 75-91. http://dx.doi.org/10.1080/1031461X.2015.1122072

Braun, V & Clarke, V. 2013. *Successful qualitative research: A practical guide for beginners*. Thousand Oaks: Sage.

Bryman, A, Bell, E, Hirschsohn, P, Dos Santos, A, Du Toit, J, Masenge, A, Van Ardt, I & Wagner, C. 2014. Research *methodology: Business management context*. Cape Town: Oxford University Press.

Bryman, A. 2016. *Social research methods.* 5th ed. London: Oxford University Press.

Brynard, DJ, Hanekom, SX & Brynard, PA. 2014. *Introduction to research.* 3rd ed. Pretoria: Van Schaik.

Busetto, L, Wick, W & Gumbinger, C. 2020. How to use and assess qualitative research methods. *Neurological Research and practice*, *2*: 1-10. https://doi.org/10.1186/s42466-020-00059-z

Butler, RP. 2008. Oral history as educational technology research. *TechTrends* 52(6): 35. https://doi.org/10.1007/s11528-008-0175-4

Cassell, C, Cunliffe, AL & Grandy, G. 2018. *The SAGE handbook of qualitative business and management research methods: History and traditions*. London: Sage. http://dx.doi.org/10.4135/9781526430212

Chancellor, R & Lee, S. 2016. Storytelling, oral history, and building the library community. *Storytelling, Self, Society* 12 (1): 39-54. https://doi.org/10.13110/storselfsoci.12.1.0039

Chapman, A. 2021. Trials of metadata: Emerging schemas for videogame cataloguing. *Journal of Library Metadata* 21(3-4): 63-103. https://doi.org/10.1080/19386389.2021.2007729

Chelliah, SL. 2023. Making photographs in language archives maximally useful: Metadata guidelines for community and academic depositors. http://dx.doi.org/10.12794/langarc2114301

Chernevych, A. 2019. "NANCY MACKAY, Curating Oral Histories: From Interview to Archive". Archivaria 87: 188-93. https://archivaria.ca/index.php/archivaria/article/view/13681 (Accessed 5 July 2023).

Cherry, A & Mukunda, K. 2015. A case study in indigenous classification: revisiting and reviving the Brian Deer scheme. *Cataloguing & Classification Quarterly* 53(5-6): 548-567. http://dx.doi.org/10.1080/01639374.2015.1008717

Chung, S & Clandinin, DJ. 2023. Narrative inquiry. In Robert, J. Tierney, Fazal Rizvi & Kadriye Ercikan (eds.), *International Encyclopedia of Education.* 4th ed. Elsevier, pp. 123-130. https://doi.org/10.1016/B978-0-12-818630-5.11014-0.

Clarke, V & Braun, V. 2013. Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist* 26(2): 120-123.

Cohen, L, Manion, L & Morrison, K. 2018. *Research methods in education*. 8th ed. New York: Routledge.

Cohen, S. 2013. Shifting questions: New paradigms for oral history in a digital world. *The Oral History Review* 40(1): 154-167. https://doi.org/10.1093/ohr/oht036

Colla, D, Goy, A, Leontino, M, Magro, D & Picardi, C. 2020. Bringing semantics into historical archive s with computer-aided rich metadata generation. *Journal on Computing and Cultural Heritage* 13: 1-36. https://doi.org/10.1145/3484398

Corrado, EM & Moulaison, HL. 2014. *Digital Preservation in Libraries, Archives, and Museums*. Lanham, MD: Rowman & Littlefield,

Corrado, EM. 2018. Discovery products and the open archives initiative protocol for metadata harvesting. *International Information & Library Review 50*(1):47-53.

Corrado, EM. 2021. Artificial intelligence: The possibilities for metadata creation. *Technical Services Quarterly* 38(4): 395-405. https://doi.org/10.1080/07317131.2021.1973797

Crawford, LM. 2019. Conceptual and theoretical frameworks in research. In Burkholder, G.J., Cox, K.A., Crawford, L.M. & Hitchcock, J., *Research design and methods: An applied guide for the scholar-practitioner*. Thousand Oaks: Sage.

Creswell, JW & Creswell, JD. 2018. *Qualitative, quantitative, and mixed methods approaches.* 5th ed. Los Angeles: Sage.

Creswell, JW & Plano Clark, VL. 2018. *Designing and conducting mixed methods research*. 3rd ed. Thousand Oaks, CA: Sage.

Creswell, JW & Poth, CN. 2016. *Qualitative inquiry and research design: Choosing among five approaches.* 4th ed. London: Sage.

Creswell, JW. 2014. *Research design: Qualitative, quantitative, and mixed methods approach*. London: Sage.

Crowe, KM & Meagher, ES. 2015. Metadata best practices for university archives images. *Journal of Library Metadata* 15(2): 79-101. https://doi.org/10.1080/19386389.2015.1041852

Da Graça Simões, M, De Freitas, MCV & Rodríguez-Bravo, B. 2016. Theory of classification and classification in libraries and archives. *Knowledge Organisation* 43(7): 530-538.

Dashrath, VB. 2014. Role of metadata in digital resource management. *International Journal of Digital Library Services* 4(3): 209-217.

DataCite Metadata Working Group. 2021. DataCite metadata schema documentation for the publication and citation of research data and other research outputs. Version 4.4. DataCite. https://doi.org/10.14454/3w3z-sa82

De La Porte, B & Higgs, R. 2019. Challenges in digitisation of cultural heritage material in the Western Cape, South Africa. *South African Journal of Information Management* 21(1):1-11. https://doi.org/10.4102/sajim.v21i1.1042

Demichev, A, Kryukov, A & Prikhodko, N. 2018. The approach to managing provenance metadata and data access rights in distributed storage using the hyperledger blockchain. https://doi.org/10.1109/ISPRAS.2018.00028

Denscombe, M. 2014. *The good research guide: for small scale social research project*. 6th ed. London: McGraw Hill Education.

Denzin, NK & Lincoln, YS. 2018. *The SAGE handbook of qualitative research*. 5th ed. Los Angeles: Sage.

Dobreski, B, Park, J, Leathers, A & Qin, J. 2020. Remodeling archival metadata descriptions for linked archives. *International Conference on Dublin Core and Metadata Applications*, 1-11. https://doi.org/10.23106/dcmi.952141553

Donaldson, D.R., Zegler-Poleska, E. and Yarmey, L., 2020. Data managers' perspectives on OAIS designated communities and the FAIR principles: mediation, tools and conceptual models. *Journal of Documentation 76*(6):1261-1277.https://doi.org/10.1108/JD-10-2019-020

Donaldson, DR, Zegler-Poleska, E & Yarmey, L. 2020. Data managers' perspectives on OAIS designated communities and the FAIR principles: mediation, tools and conceptual models. *Journal of Documentation 76*(6):1261-1277.https://doi.org/10.1108/JD-10-2019-020

Don-Solomon, A & Eke, GT. 2018. Ontological & epistemological philosophies underlying theory building: a scholarly dilemma or axiomatic illumination – the business research perspective. *Journal of Business and Innovation Research* 6(2): 1-7.

Dooley, JM, Farrell, KS, Kim, T & Venlet, J. 2017. Developing web archiving metadata best practices to meet user needs. *Journal of Western Archives* 8(2). https://doi.org/10.26077/cffd-294a

Dooley, M, Moore, E & Vallejo, C. 2017. Research ethics. In *Moore, E. & Dooly, M. (eds.), Qualitative approaches to research on plurilingual education (pp. 351-362)*. https://doi.org/10.14705/rpnet.2017.emmd 2016.63

Dougherty, J & Simpson, C. 2012. Who owns oral history? a creative commons solution. In D. Boyd, S. Cohen, B. Rakerd, & D. Rehberger (Eds.), Oral history in the digital age. Available at: http://ohda.matrix.msu.edu/2012/06/a-creative-commons-solution/. (Accessed 2 November 2023)

Du Plooy-Cilliers, F, Davis, C & Bezuidenhout, RM. (eds). 2014. *Research matters*. Cape Town: Juta.

Dube, L, Mhlongo, M & Ngulube, P. 2014. The ethics of anonymity and confidentiality: Reading from the University of South Africa Policy on Research Ethics *Indilinga – African Journal of Indigenous Knowledge Systems* 13(2): 201-212.

Dudovskiy, J. 2018. The ultimate guide to writing a dissertation in business studies: A step-by-step assistance. BRM Edition. Available at: https://research-methodology.net/about-us/ebook/ (Accessed 2 November 2023)

Duranti, L, Rogers, C, Mcleod, J & Shepherd, E. 2019. *Trusting records and data in the cloud: The creation, management, and preservation of trustworthy digital content*. London: Facet.

Duranti, L. 1994. The records: Where archival universality resides. *Archival Issues* 19(2): 83-94. https://www.jstor.org/stable/41101882(Accessed 16 December 2021).

Duval, E, Hodgins, W, Sutton, S & Weibel, SL. 2002. Metadata principles and practicalities. *D-lib Magazine* 8(4): 1-10.

Ebohon, OJ, Ajayi, TO & Ganiyu, SA. 2021. Understanding research paradigm in social sciences: A critique of two papers on critical success factors for BIM Implementation. *Journal of Environmental Technology* 2: 64-70.

Edmondson, R. 2016. *Audiovisual archiving: philosophy and principles*. 3rd ed. Paris: UNESCO

Eíto-Brun, R. 2015. A metadata infrastructure for a repository of civil engineering records: EAC-CPF as a cornerstone for content publishing. *Journal of Archival Organization* 12(1-2): 62-76. https://doi.org/10.1080/15332748.2015.999502

Elkind, D. 2004. The problem with constructivism. *The Educational Forum,* 68: 306-312. https://doi.org/10.1080/00131720408984646

Farnel, S, Shiri, A, Campbell, S, Cockney, C, Rathi, D & Stobbs, R. 2017. A community-driven metadata framework for describing cultural resources: The Digital Library North project. *Cataloguing & Classification Quarterly* 55(5): 289-306. https://doi.org/10.1080/01639374.2017.1312723.

Feliciati, P & Alfier, A. 2013. Archives on the web and users expectations: Towards a convergence with digital libraries. Rev. Nat. Cent. Digitization 22:81-92 http://elib.mi.sanu.ac.rs/files/journals/ncd/22/ncd22081.pdf (Accessed 16 December 2021).

Feliciati, P & Alfier, A. 2014. From access to use: Premises for a user-centered quality model for the development of archives online. In Bolikowski, Ł, Casarosa, V, Goodale, P, Houssos, N, Manghi, P. and Schirrwagen, J. eds. *Theory and Practice of Digital Libraries* 416:174 179. https://doi.org/10.1007/978-3-319-08425-1_18

Feliciati, P. 2023. "Archival Users and AI Tools for Reference and Access: A Study Within the InterPARES Trust AI Project*". Italian Journal of Library, Archives and Information Science (JIS*.It 14 (3):117-28. https://doi.org/10.36253/jlis.it-558.

Fell, S & Ou, C. 2022. Revisiting Oral History Cataloging Policies for Multi-Format Discovery in MARC: A Case Study. *Technical Services Quarterly 39*(3): 225-240. https://doi.org/10.1080/07317131.2022.2082655

Flick, U. 2018. *An introduction to qualitative research.* 6th ed. Los Angeles: Sage.

Flick, U. 2022. *The SAGE handbook of qualitative Research Design*. Los Angeles: Sage.

Formenton, D & De Souza Gracioso, L. 2022. Metadata standards in web archiving technological resources for ensuring the digital preservation of archived websites. *Digital Journal of Library and Information Science* 20:e022001. Available at: https://www.researchgate.net/publication/357780359 (Accessed 26 December 2023).

Fouka, G & Mantzorou, M. 2011. What are the major ethical issues in conducting research? Is there a conflict between the research ethics and the nature of nursing? *Health Science Journal* 5: 3-14.

Franks, PC. 2013. *Records and information management*. Chicago: American Library Association.

Fraser, I & Halvey, M. 2017. On designing an oral history search system. *Journal of Documentation* 73(5): 1-16.

Frisch, M. 2016. Oral history in the digital age: Beyond the raw and the cooked. *Australian Historical Studies* 47(1): 92-107. https://doi.org/10.1080/1031461X.2015.1122073

Gajjar, NB. 2013. Ethical consideration in research. *International Journal for Research in Education*, 2(7). Accessible at: http://www.raijmr.com/ijre/wp-content/uploads/2017/11/IJRE_2013_vol02_issue_07_02.pdf (Accessed 16 December 2021).

Gall, MD, Gall, JP & Borg, WR. 2003. *Educational research: An introduction*. 7th ed. Boston, MA: Pearson.

Garcia, AJ. 2018. Social implications of research. In M. Allen (ed.), *The SAGE Encyclopedia of communication research methods*. Sage. https://doi.org/10.4135/9781483381411.

Gartner, R. 2016. *Metadata: Shaping knowledge from antiquity to the semantic web*. Cham, Switzerland: Springer.

Gibbs, G. 2007. *The Sage qualitative research kit: Analysing qualitative data.* Sage Publications Ltd. https://doi.org/10.4135/9781849208574

Gill, H, Vos, J, Villa-Torres, L & Ramirez, MS. 2019. Migration and inclusive transnational heritage: Digital innovation and the new roots Latino Oral History Initiative. *The Oral History Review* 46(2): 277-299. https://doi.org/10.1093/ohr/ohz011

Gilliland, A & Willer, M. 2014. Metadata for the information multiverse. *In Proceedings of iConference 2014*, Berlin, Germany, 4-7 March 2014.

Gilliland, AJ. 2000. *Setting the stage. Introduction to metadata: Pathways to digital information,* Getty Research Institute. http://www.getty.edu/research/institute/standards/intrometadata/ (Accessed 16 December 2021).

Gilliland-Swetland, AJ. 2000. *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment*. Massachusetts: Council on Library and Information Resources. Available at: https://files.eric.ed.gov/fulltext/ED440660.pdf (Accessed 16 August 2021).

Giovanni, C, Blanke, T, Jeurgens, C & Noordegraaf, J. 2021. Archives and AI: An overview of current debates and future perspectives. *Journal on Computing and Cultural Heritage* 15(1). https://doi.org/10.1145/3479010

Goh, EMY. 2016. *Archival law from the trenches : the impact of archival legislation on records management in commonwealth countries* (T). University of British Columbia. https://dx.doi.org/10.14288/1.0300019

Goodchild, B, Ambrose, A & Maye-Banbury, A. 2017. Storytelling as oral history: revealing the changing experience of home heating in England. *Energy Research & Social Science* 31:137-144. https://doi.org/10.1016/j.erss.2017.06.009

Gracy, KF & Lambert, F. 2014. Who's ready to surf the next wave? A study of perceived challenges to implementing new and revised standards for archival description. *The American Archivist* 77(1): 96-132. https://doi.org/10.17723/aarc.77.1.b241071w5r252612

Grant, C & Osanloo, A. 2014. Understanding, selecting, and integrating a theoretical framework in dissertation research: creating the blueprint for your "house". *Administrative Issues Journal: Connecting Education, Practice and Research*, 40(2): 12-26.

Grassi, PA, Lefkovitz, NB, Nadeau, EM, Galluzzo, RJ, Dinh, AT & Rosslyn, VA. 2018. *Attribute metadata: A proposed schema for evaluating federated attributes*. National Institute of Standards and Technology Internal Report 8112, US Department. https://doi.org/10.6028/NIST.IR.8112

Greenberg, J. 2005. Understanding metadata and metadata schemes. *Cataloging & Classification Quarterly* 40(3-4): 17-36.

Greenberg, J. 2017. Big metadata, smart metadata, and metadata capital: Toward greater synergy between data science and metadata. *Journal of Data and Information* Science 2(3). https://doi.org/10.1515/jdis-2017-0012

Grimsley, RL & Wynne, SC. 2009. Creating access to oral histories in academic libraries. *College & Undergraduate Libraries* 16(4): 278-299. https://doi.org/10.1080/10691310903355903

Groenewald, R & Breytenbach, A. 2011. The use of metadata and preservation methods for continuous access to digital data. *Electronic Library* 29(2): 236-248. http://hdl.handle.net/2263/17635 (Accessed 16 December 2021).

Guetzkow, J, Lamont, M & Mallard, G. 2004. What is originality in the humanities and the social sciences? *American Sociological Review,* 69(2): 190-212. https://doi.org/10.1177%2F000312240406900203

Halamy, S & Kibat, SA. 2017. Reconnecting the past through oral history: An effort in preserving Malaysians' intangible cultural heritage. In: Saufi, A., Andilolo, I., Othman,

N. & Lew, A. (eds). Balancing development and sustainability in tourism destinations. Springer, Singapore. https://doi.org/10.1007/978-981-10-1718-6_4

Hammond, M & Wellington, J. 2013. *Research methods: The key concepts*. Abingdon: Routledge.

Hancock, DR & Algozinne, B. 2017. *Doing case study research: A practical guide for beginning researchers*. New York: Teachers College Press.

Harrison, H, Birks, M, Franklin, R, & Mills, J. 2017. Case Study Research: Foundations and Methodological Orientations. Forum Qualitative Sozialforschung Forum: Qualitative Social Research, 18(1). https://doi.org/10.17169/fqs-18.1.2655

Hart, C. 1998. *Doing a literature review: Releasing the social science research imagination*. London: Sage.

Hart, TR & De Vries, D. 2017. *Metadata provenance and vulnerability. Information technology and libraries*. https://doi.org/10.6017/ital.v36i4.10146

Haynes, D. 2018. Metadata for Information Management and Retrieval: understanding metadata and its use. London: Facet Publishing.

Hendrich, G. 2017. A rich storehouse for research: The historical development of the Western Cape archives and Records service. *Journal for Contemporary History* 42(2): 74-97. https://dx.doi.org/10.18820/24150509/JCH42.v2.4

Hendrigan, H. 2016. An examination of oral history and archival practices among graduate students in select Canadian Comprehensive Research Universities. *Oral History Forum d'histoire Orale* 36. Available at:http://www.oralhistoryforum.ca/index.php/ohf/article/view/614 (Accessed 16 October 2021).

Herawan, L & Sensuse, DI. 2018. Analysis of Factors for Adoption and Use of AtoM in Indonesia. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 436-441). IEEE. https://doi.org/10.1109/ICITSI.2018.8696086

Hider, P. 2015. A survey of the coverage and methodologies of schemas and vocabularies used to describe information resources. *Knowledge Organisation,* 42(3): 154. https://doi.org/10.5771/0943-7444-2015-3-154 .

Hu, X, Ng, TJ & Xia, S. 2018. User-centered evaluation of metadata schema for nonmovable cultural heritage: Murals and stone cave temples. *Journal of the Association for Information Science and Technology* 69(12): 1476-1487. https://doi.org/10.1002/asi.24065

International Association of Sound and audiovisual Archives. 2021. https://www.iasa-web.org/ (Accessed 20 October 2021)

International Council of Archives (ICA). 2016. https://www.ica.org/ (Accessed 23 January 2022)

International Organization for Standardization. 2016. ISO 15489-1 Information and Documentation-Records Management. Part 1: Concepts and Principles. Available at: https://www.iso.org/standard/62542.html. (Accessed 1 November 2021)

Jacobs, L. 2020. Embedding ethical principles in the information science research process. In P. Ngulube (ed). *Handbook of research on connecting research methods for information science.* Hershey: IGI Global, 52-79. https://doi.org/10.4018/978-1-7998-1471-9.ch004

Jaillant, L & Caputo, A. 2022. Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data. *AI & Society* 37(3): 823-835. https://doi.org/10.1007/s00146-021-01367-x

Jana, A & Rout, R. 2022. Designing metadata schema for a human library: a prototype. *Digital Library Perspectives 38*(3): 346-361. http://dx.doi.org/10.1108/DLP-07-2021-0052

Jason, L & Glenwick, D. (eds). 2016. *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods*. New York: Oxford University Press.

Johnson, R & Christensen, L. 2014. *Educational research quantitative, qualitative, and mixed approaches*. 5th ed. London: Sage.

Kapidakis, S. 2018. Metadata synthesis and updates on collections harvested using the open archive initiative protocol for metadata harvesting. In *Digital Libraries for Open Knowledge: 22nd International Conference on Theory and Practice of Digital Libraries, TPDL 2018, Porto, Portugal, September 10–13, 2018, Proceedings 22* (pp. 16-31). Springer International Publishing. https://doi.org/10.1007/978-3-030-00066-0_2

Kari, KH & Baro, EE. 2016. Digital preservation practices in university libraries: A survey of institutional repositories in Nigeria. *Preservation, Digital Technology & Culture* 45(3):134-144. https://doi.org/10.1515/pdtc-2016-0006

Kata, L, Milbrodt, N, Sielaff, S & Vos, J. 2020. Oral history metadata and description: A survey of practices. *TX: Oral History Association*: Oral History Association. Available at: https://www.oralhistory.org/wp-content/uploads/2021/01/OHA-MTF-White-Paper_2020.pdf (Accessed,22 July 2023).

Katuu, S & Van der Walt, T. 2016. Assessing the legislative and regulatory framework supporting the management of records in South Africa's public health sector. *South African Journal of Information Management 18*(1):1-7.

Katuu, S. 2023. Tapestry of the education and training landscape for archives and records management in Africa. In Mpho Ngoepe (ed) *Managing Digital Records in Africa*. New York: Routledge.

Kelly, JM & Horan, J. 2021. Archive as pedagogy: Oral history and a journal of the plague year. *Collections: A Journal for Museum and Archives Professionals* 17(3): 237-246. https://doi.org/10.1177/1550190620981029

Kenneavy, K & Harnos, CE. 2023. *Social research methods: Sociology in Action.* Thousand Oaks, CA: Sage.

Kenneavy, K, Harnois, CE, Atkinson, MP and Korgen, KO (eds.). 2022. *Social research methods: Sociology in action*. London: Sage.

Kennedy, MR. 2008. Nine questions to guide you in choosing a metadata schema. *Journal of digital Information* 9(1). Available at: https://jodi-ojs-tdl.tdl.org/jodi/article/view/226 (Accessed 16 December 2021).

Khaldi, K. 2017. Quantitative, qualitative or mixed research: Which research paradigm to use? *Journal of Educational and Social Research* 7(2): 15-24.

Kim, J & Liu, M. (eds.). 2017. *The SAGE encyclopedia of communication research methods*. Vol. 4. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781483381411

Kim, JY. 2020. Protecting metadata of access indicator and region of interests for image files. *Security and Communication Networks*. https://doi.org/10.1155/2020/4836109

Király, P. 2019. *Measuring metadata quality*. Thesis Doctoral Degree in Philosophy at the Faculty of Humanities of the Georg-August-Universität Göttingen. Göttingen, Germany.

Kivunja, C & Kuyini, AB. 2017. Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education* 6(5): 26-41.

Knight, FT. 2022.Term circles: Using linked data as a tool to mitigate colonial subject bias. *Journal of Library Metadata* 22(1-2): 105-133. https://doi.org/10.1080/19386389.2022.2051980

Knight, RC, Rodrigues, E & Ciota,R. 2020. Facilitating collaborative metadata creation for faculty-initiated digital projects. *Journal of Library Metadata,* 51. https://doi.org/10.1080/19386389.2020.1728479

Kour, S. 2014. Ethical and legal issues in educational research. *Indian Journal of Applied Research* 4(6): 133-135.

Kumar, R. 2011. *Research methodology: A step by step guide for beginners*. London: Sage.

Kumar, R. 2014. *Research methodology: A step-by-step guide for beginners*. 4th ed. London: Sage.

Kumatongo, B & Muzata, KK. 2021. Research paradigms and designs with their application in education. *Journal of Lexicography and Terminology 5*(1):16-32.

Ladley, J. 2012. *Data governance: How to design, deploy, and sustain an effective data governance program*. Waltham: Morgan Kaufmann.

Laher, S, Fynn, A & Kramer, S. (eds). 2019. *Transforming research methods in the social sciences: Case studies from South Africa*. Johannesburg: Wits University Press.

Laher, S. 2016. Ostinato rigore: Establishing methodological rigour in quantitative research. *South African Journal of Psychology* 46(3): 316-327. https://doi.org/10.1177/0081246316649121

Lambert, D & Frisch, M. 2013. Digital curation through information cartography: A commentary on oral history in the digital age from a content management point of view. *Oral History Review* 40(1): 135-153. https://doi.org/10.1093/ohr/oht035.

Lapp, J.M. 2023."The only way we knew how:" provenancial fabulation in archives of feminist materials. *Arch Sci 23*, 117–136 (2023). https://doi.org/10.1007/s10502-021-09376-x.

Larson, RR & Janakiraman, K. 2019. Connecting archival collections: The Social Networks and Archival Context Project. In *Research and Advanced Technology for Digital Libraries* (pp. 3-14). *Heidelberg, Dordrecht*. London, New York: Springer. https://doi.org/10.1007/978-3-642-24469-8_3.

Lavoie, BF. 2000. Meeting the challenges of digital preservation: The OAIS reference model. Available at: The OAIS reference model (oclc.org). (Accessed 29 April 2024).

Lee, E. 2017. Knowledge was their treasure: Applying KO approaches to archaeological research. *Knowledge Organisation* 44(8): 644-655.

Leedy, PD & Ormrod, JE. 2015. *Practical research: Planning and design.* 11th ed. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Leedy, PD. & Ormrod, JE. 2013. *Practical research: Planning and design.* 10th ed. New Jersey: Pearson Education.

Lincoln, YS, Lynham, SA & Guba, EG. 2011. Paradigmatic controversies, contradiction, and emerging confluence. In Denzin, N. & Lincoln, Y.S. (eds.), *The Sage handbook of qualitative research*. Thousand Oaks: Sage.

Litchman, M. 2014. *Qualitative research for the social science*. Los Angeles: Sage.

Loshin, D. 2013. Metadata. In *MK Series on Business Intelligence, Business Intelligence.* 2nd ed. Morgan Kaufmann, pp. 119-130. https://doi.org/10.1016/B978-0-12-385889-4.00009-0.

Lune, H & Berg, BL. 2017. *Qualitative research methods for the social sciences*. 9th ed. Boston: Pearson Education.

MacKay, N. 2016. *Curating oral histories: From interview to archive*. New York:Routledge.

Macleod, C & Mnyaka, P. 2018. Introduction: The politics of anonymity and confidentiality. In *The Palgrave handbook of ethics in critical research* (pp. 227-240) https://doi.org/10.1007/978-3-319-74721-7_15

Maguire, M & Delahunt, B. 2017. Doing thematic analysis: A practical, step-by-step guide for teaching and learning scholars. *All-Ireland Journal of Teaching and Learning in Higher Education* 3: 351-314.

Maluleka, JR, Nkwe, M & Ngoepe, MP. 2018. Provision for digital records management in the curriculum of LIS schools at public universities in South Africa. *Journal of South African Society of Archivists* 51.

Maree, K. 2017. *First steps in research.* 2nd ed. Pretoria: Van Schaik.

Maria, MS, Cruz, CE, San Juan, MJS, Tan, CY & Amarnani, R. 2022. The concept of originality and creativity in the humanities, the social sciences and the natural sciences. *Asia Pacific Journal of IP Management and Innovation,* 1: 1.

Masenya, TM & Ngulube, P. 2021. Digital preservation systems and technologies in South African academic libraries. *South African Journal of Information Management* 23(1): 1249. https://doi.org/10.4102/sajim.v23i1.1249

Masenya, TM. 2018. *A framework for preservation of digital resources in academic libraries in South Africa* (Doctoral dissertation, PhD thesis, University of South Africa).

Matthews, B & Ross, L. 2010. *Research methods*. London: Pearson Longman.

Matusiak, KK, Tyler, A, Newton, C & Polepeddi, P. 2017. Finding access and digital preservation solutions for a digitized oral history project: A case study". *Digital Library Perspectives* 33(2): 88-99. https://doi.org/10.1108/DLP-07-2016-0025

Matusiak, KK, Werling, S, Donovan, L & Carlson, S. 2021. Rights metadata in a community archive: Implementing standardized rights statements. *Journal of Library Metadata* 21(1-2): 27-44. https://doi.org/10.1080/19386389.2021.1955586

Maxwell, J. 2013. *Qualitative research design: An interactive approach*. Thousand Oaks: CA: Sage.

Mazé, E. 2014. Deconstruction without destruction: Creating metadata for oral history in a digital world. In Boyd, D.A. & Larson, M.A. (eds.), *Oral History and Digital Humanities* (pp. 141-162). Palgrave Macmillan. https://doi.org/10.1057/9781137322029_9

McCartney, L. 2017. The evolution and issues encountered in providing meaningful and useful online access to oral history recordings. *International Association of Sound and Audiovisual Archives (IASA Journal*)47: 30-40.

McDonald, M, Puddicombe, A, Weymouth, A & Williams, A. 2021. Methods for Collecting Oral Histories Amidst a Crisis. *The Serials Librarian 81*(2):120-131. https://doi.org/10.1080/0361526X.2021.1985685

Merriam, SB. 2009. Qualitative research: a guide to design and implementation. San Francisco: Jossey-Bass.

Meriam Webster Dictionary. https://www.merriam-webster.com/ (Accessed 29 April 2022).

Miles, MB, Huberman, AM & Saldana, J. 2014. *Qualitative data analysis: A methods sourcebook*. 3rd ed. Los Angeles: Sage.

Miller, SJ. 2022. *Metadata for digital collections: A how-to-do-it manual*. 2nd ed. New York: Neal-Schuman.

Molefe, N & Schellnack-Kelly, I. 2019. Users' perceptions of services at the National Archives and Records Service of South Africa (NARSSA). *Innovation: journal of appropriate librarianship and information work in Southern Africa 2019*(58): 3-15.

Molobye, K. 2014. Building a new home of our history – the Bafokeng Digital Archive. *Journal of the South African Society of Archivists* 47: 50-60.

Moon, K & Blackman, D. 2014. A Guide to understanding social science research for natural scientists. *Conservation Biology*, 28: 1167-1177. Available at: http://onlinelibrary.wiley.com/doi/10.1111/cobi.12326/full

Morgan, DL. 2019. After triangulation, what next? *Journal of Mixed Methods Research,* 13(1): 6-14. https://doi.org/10.1177/1558689818780596

Mosako, DR & Ngoepe, M. 2021. New ways to access archival and special collections. *Library Hi Tech News* 38(3): 18-20. https://doi.org/10.1108/LHTN-05-2021-113

Mosako, DR. 2022. Implementation of museum educational strategies through information communication technology: Building on the Global Sustainable Development Goal 4. *South African Museums Association Bulletin* 44(1): 29-37.

Mosako, DR. 2023. Dealing with museum educational tours as heritage knowledge sharing practices in South Africa during the COVID-19 pandemic. In Delores V Mullings, Olasumbo Adelakun and Jennifer Clarke(eds). *African, Caribbean and Black People's Resilience During COVID-19.*Canada: Demeter Press.

Moser, A & Korstjens, I. 2018. Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. *European Journal of General Practice* 24(1): 9-18. https://doi.org/10.1080/13814788.2017.1375091

Moses, JW & Knutsen, TL. 2019. *Way of knowing: Competing methodologies in social and political research*. 3rd ed. London: Red Globe Press

Moss, WW & Mazikana, PC. 1986. *Archives, oral history and oral tradition: A RAMP study.* Paris: UNESCO.

Mosweu, O. 2011. Performance audit in the Botswana public service and arising records management issues. *Journal of the South African Society of Archivists* 44: 107-115.

Mukwevho, J & Ngoepe, M. 2019. Taking archives to the people: The use of social media as a tool to promote public archives in South Africa. *Library Hi Tech 37*(3): 374-388. https://doi.org/10.1108/LHT-11-2017-0228

Mukwevho, J & Ngulube, P. 2022. A framework for enhancing the visibility and accessibility of public archives in South Africa. *Archives and Records: The Journal of the Archives and Records Association* 42(3): 1-9. https://doi.org/10.1080/232 57962.2022.2031933

Murillo, AP. 2014. Data at risk initiative: examining and facilitating the scientific process in relation to endangered data. *Data Science Journal*, *12:* 207-219. https://doi.org/10.2481/dsj.12-048

National Archives and Records Service of South Africa. 2006. *Managing electronic records in governmental bodies: Metadata requirements.* NARS

National Archives and Records Services of Sout Africa. 2016. https://nationalgovernment.co.za/units/view/123/national-archives-and-records-service-of-south-africa-narssa

National Archives and Records Srevices of Sout Africa Act No.43 of 1996. https://www.gov.za/documents/national-archives-south-africa-act (Accessed 20 October 2021)

Ncala, BNM. 2017. *Preservation of, and access to audio-visual records at the National Film, Video and Sound Archives of South Africa*. Master's thesis, University of KwaZulu-Natal, South Africa.

Nelson, ML.2001. Better interoperability through the Open Archives Initiative. *New Review of Information Networking* 7(1): 133–145. https://doi.org/10.1080/13614570109516973

Netshakhuma, NS. 2019. The role of archives and records management legislation after colonialism in Africa: Case of Southern Africa. *Records management Journal* 29 (1/2: 21 223. https://doi.org/10.1108/RMJ-09-2018-0024

Neuman, WL. 2011. *Social research methods: Qualitative and quantitative approaches*. 6th ed. Boston: Pearson Education.

Ngoasheng, C, Ngoepe, M & Marutha, NS. 2021. Sounds like a broken record: Preservation and access of audio-visual records at the South African broadcasting corporation radio. *Global Knowledge, Memory and Communication.* https://doi.org/10.1108/GKMC-10-2020-0153

Ngoepe, M & Katuu. 2017. Provision of records created in networked environments in the curricula of institutions of higher learning in Africa. *New Review of Information Networking* 22(1):1-12. https://doi.org/10.1080/13614576.2017.1297731

Ngoepe, M & Keakopa, S. 2011. An assessment of the people in South Africa: Can we really build the Tower of Babel. *ESARBICA* -Botswana Comparison. *Records Management Journal* 21(2): 122-134. https://doi.org/10.1108/09565691111152071

Ngoepe, M & Ngulube, P. 2011. Assessing the extent to which the National Archives and Records Service of South Africa has fulfilled its mandate of taking the archives to the people. *Innovation: Journal of Appropriate Librarianship and Information Work in Southern Africa* 42: 3-22.

Ngoepe, M & Saurombe, N. 2021. Africanisation of the South African archival curriculum: A preliminary study of undergraduate courses in an open distance e-learning environment. *Education for Information 37*(1): 53-68.

Ngoepe, M. 2019. Archives without archives: A window of opportunity to build inclusive archive in South Africa. *Journal of the South African Society of Archivists* 52: 149-166. https://orcid.org/0000-0002-6241-161X

Ngoepe, M. 2020. Whose truth is true? the use of archival principles to authenticate oral history. In P. Ngulube (ed). *Handbook of Research on Connecting Research Methods for Information Science*. Hershey: IGI Global, 307-319. https://doi.org/10.4018/978-1-7998-1471-9.ch016

Ngoepe, M. 2022. Archiving by state has regressed in post-apartheid SA: Digitalisation will also ensure records are accessible. Will Also Ensure Records Are Accessible. *Sowetan*. Available at: https://www.sowetanlive.co.za/opinion/columnists/2022-07-14-mpho-ngoepe-archiving-by-state-has-regressed-in-postapartheid-sa/(Accessed 11 December 2022).

Ngulube, P, Sibanda, P & Makoni, NLS. 2013. Mapping access and use of archival materials held at the Bulawayo Archives in Zimbabwe. *ESARBICA Journal,* 32: 123-137. https://doi.org/10.4314/esarjo.v32i0.

Ngulube, P. 2015. Qualitative data analysis and interpretation: systematic search for meaning. In Mathipa, E.R. & Gumbo, M.T. (eds), *Addressing research challenges: Making headway for developing researchers*. Mosala-MASEDI Publishers & Booksellers: Noordywk, pp. 131-156.

Ngulube, P. 2018. Overcoming the difficulties associated with using conceptual and theoretical frameworks in heritage studies. In Ngulube, P. (Ed.), *Handbook of research on heritage management and preservation*. Hershey, PA: IGI Global, pp. 1-23. https://doi.org/10.4018/978-1-5225-3137-1.ch001

Ngulube, P. 2020. Theory and theorising in information science scholarship. In Ngulube, P. (ed.), *Handbook of research on connecting research methods for information science research*. Hershey, PA: IGI Global, pp. 18-39. https://doi.org/10.4018/978-1-7998-1471-9.ch002

Nimer, C. 2010. RDA and archives. *Journal of Archival Organization* 8(3-4): 239.

Niu, J. 2013. Recordkeeping metadata and archival description: A revisit. *Archives and Manuscripts* 41(3): 203-215. https://doi.org/10.1080/01576895.2013.829751

Noor, MU. 2020. Description methods for image records using Access to Memory (AtoM) system for better records retrieval. Advances in social science, education and humanities

research. *Proceedings of International Conference on Vocational Higher Education* 426:73-78. https://doi.org/10.2991/assehr.k.200331.169

Note, M. 2020. The effect of technology on archival activities. Available at: https://lucidea.com/blog/the-effect-of-technology-on-archival-activities/ (Accessed 16 December 2021).

Nowell, LS, Norris, JM, White, DE & Moules, NJ. 2017. Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods* 16(1): 1-13. https://doi.org/10.1177/1609406917733847

Nsibirwa, Z. 2019. The Ethical implications of digitisation of the Alan Paton Centre Struggle Archives. *ESARBICA Journal* 38: 83-106. https://doi.org/10.4314/esarjo.v38i1.5

Ocholla, DN. 1999. Insight into information seeking and communication behaviour of academics. *International Information and Library Review* 31(3): 119-143. https://doi.org/10.1006/iilr.2000.0116

Oelofse, M & Du Bruyn, D. 2004. The importance of oral history in a transforming South Africa. *Journal for Contemporary History 29*(1): 151-169.

Openarchives. 2012.The Open Archive Initiative Protocol for Metadata Harvesting. Available at : https://www.openarchives.org/OAI/openarchivesprotocol.html (Accessed 26April *2024*

Opgenhaffen, L. 2022. Archives in action. The impact of digital technology on archaeological recording strategies and ensuing open research archives. *Digital Applications in Archaeology and Cultural Heritage* 27: e00231. https://doi.org/10.1016/j.daach.2022.e00231

Oral History Association (OHA). 2019. "*Oral history best practices*." https://www.oralhistory.org/wp-content/uploads/2018/07/OHA-best-practices-2018.pdf (accessed 1 August 2023).

Palavitsinis, N, Manouselis, N & Salvador, S. 2017. Metadata and quality in digital repositories and libraries from 1995 to 2015: A literature analysis and classification. *International Information & Library Review* 49(3): 176-186. https://doi.org/10.1080/10572317.2016.1278194

Park, J & Tosaka, Y. 2010. Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. *Information Technology & Libraries* 29(3): 104-116. https://doi.org/10.6017/ital.v29i3.3136

Patterton, LH, Bothma, TJ & van Deventer, MJ, 2024. Risky Business: Data-At-Risk in a Dynamic and Evolving Multidisciplinary Research Environment. *Data Science Journal* 23 (11): 1-21. https://doi.org/10.5334/dsj-2024-011

Patterton, LH. 2023. Data Rescue: Defining a Comprehensive Workflow that includes the Roles and Responsibilities of the Research Library. Thesis (PhD), University of Pretoria, Pretoria. Available at: https://repository.up.ac.za/handle/2263/89829. (accessed 29 April 2

Pearce-Moses, R. 2005. *Glossary of archival and records terminology (Archival Fundamentals Series II).* Chicago: The Society of American Archivists. (Accessed 27 October 2021)

Pervin, N & Mokhtar, M. 2022. The interpretivist research paradigm: A subjective notion of a social context. *International Journal of Academic Research in Progressive Education and Development* 11(2): 419-428. http://dx.doi.org/10.6007/IJARPED/v11-i2/12938

Phillips, EM & Pugh, DS. 2005. *How to get a PhD: A handbook for students and their supervisors.* New York: McGraw-Hill Education.

Phillips, M. 2015. Metadata quality, completeness, and minimally viable records metadata quality, completeness, and minimally viable records. *Mark E. Phillips Journal.* http://vphill.com/journal/post/4075/ (Accessed 2 October 2021).

Pickard, AJ. 2013. *Research methods in information*. 2nd ed. London: Facet Publishing.

Plotkin, D. 2021. *Data stewardship: An actionable guide to effective data management and data governance*. London: Academic press.

Polit, DF & Beck, CT. 2010. *Essentials of nursing research: Appraising evidence for nursing practice.* 7th ed. Philadelphia: Lippincott Williams & Wilkins: Wolters Kluwer Health.

Polit, DF & Beck, CT. 2014. *Essentials of nursing research: Appraising evidence for nursing practice.* Lippincott Williams & Wilkins.

Poynton, C. 2012. *Digital video and HD algorithms and interfaces*. Waltam, USA: Morgan Kaufman

Pryse, J, Harp, M, Mannheimer, S, Marsolek, W & Cowles, W. 2021. *Oral history interviews data curation primer. Data curation network.* Available at: https://hdl.handle.net/11299/219052 (Accessed 20 July 2022).

Punch, K. 2014. *Introduction to social research: quantitative and qualitative approaches*. Los Angeles, California: Sage.

Raghavan, K & Krishnamurthy, M. 2013. An overview of classification technologies. *DESIDOC Journal of Library & Information Technology* 33(4): 306-313.

Ramesh, P, Vivekavardhan, J & Bharathi, K. 2015. Metadata diversity, interoperability and resource discovery issues and challenges. *DESIDOC Journal of Library & Information Technology* 35(3): 193-199. 10.14429/djlit.35.3.8074

Rawson, KJ. 2018. The rhetorical power of archival description: Classifying images of gender transgression. *Rhetoric Society Quarterly* 48(4): 327-351. https://doi.org/10.1080/02773945.2017.1347951

Rayan, R, Shimizu, C, Sieverding, H, & Hitzler, P. 2023. A Modular Ontology for MODS – Metadata Object Description Schema. In: Ortiz-Rodriguez, F., Villazón-Terrazas, B., Tiwari, S., Bobed, C. (eds) Knowledge Graphs and Semantic Web. KGSWC 2023. *Lecture Notes in Computer Science* 14382. Springer, Cham. https://doi.org/10.1007/978-3-031-47745-4_13

Reeves, A. 2013. *Managing data in motion: data integration best practice techniques and technologies*. Amsterdam: Newnes.

Rehman, AA & Khalid, A. 2016. An introduction to research paradigms. *International Journal of Educational Investigations* 3(8): 51-59. Available at: www.ijeionline.com (Accessed 23 August 2021)

Remler, DK & Van Ryzin, GG. 2011. *Research methods in practice: Strategies for description and causation*. London: Sage.

Ribeiro, F. 2014. The use of classification in archives as a means of organization, representation and retrieval of information. *KO Knowledge Organization* 41(4): 319-326.

Riley, J. 2017. Understanding metadata. *Washington DC, United States: National Information Standards Organization* 23: 7-10. *http://www.NISO.org/publications/press/UnderstandingMetadata.pdf*

Ritchie, D. 2011. Introduction: The evolution of oral history. In D. Ritchie (ed.), *The Oxford Handbook of Oral History* (pp. 12). Oxford University Press.

Ritchie, DA. 2014. *Doing oral history*.? New York: Oxford University Press.

Robinson, OC. 2014. Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology* 11(1): 25-41. https://doi.org/10.1080/14780887.2013.801543

Robson, D, Sassen, C, Thomale, J & Yanowski, K. 2019. Enhancing the discovery of tabletop games. *Library Resources & Technical Services 63*(3): 199. https://doi.org/10.5860/lrts.63n3.199

Ronzino, P, Amico, N & Niccolucci, F. 2011. *Assessment and comparison of metadata schemas for architectural heritage*. Available at: https://www.researchgate.net/publication/259780242 (Accessed 20 January 2022)

Rose, J. & Johnson, CW. 2020. Contextualizing reliability and validity in qualitative research: toward more rigorous and trustworthy qualitative social science in leisure research, *Journal of Leisure Research* 51(4): 432-451. doi:10.1080/00222216.2020.1722042

Ross, S & Hedstrom, M. 2005. Preservation research and sustainable digital libraries. *International journal on digital libraries 5*: 317-324. https://doi.org/10.1007/s00799-004-0099-3

Ryan, M, Keating, D & Finegan, J. 2022. Managing and accessing web archives: Irish practitioners' perspectives. *AI & Society* 37(3): 975-984. https://doi.org/10.1007/s00146-021-01364-0

Sambo, AS, Urhefe, EA & Ejitagha, S. 2017. A survey of digital preservation challenges in Nigerian libraries: librarians' perspective. *International Journal of Curation* 12(1):117-128. https://doi.org/10.2218/ijdc.v12i1.426

Sandaver, H. 2020. Technology, metadata and Mannix Library special collections: Descriptive tools and considerations. *The Australian and New Zealand Theological Library Association (ANZTLA) EJournal* 25: 25-46.

Sanjari, M, Bahramnezhad, F, Fomani, FK, Shoghi, M & Cheraghi, MA. 2014. Ethical challenges of researchers in qualitative studies: The necessity to develop a specific guideline. *Journal of Medical Ethics and History of Medicine* 7. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4263394/ (Accessed 31 August 2021)

Saunders, B, Kitzinger, J & Kitzinger, C. 2015. Anonymising interview data: Challenges and compromise in practice. *Qualitative research* 15(5): 616-632.

Saunders, MNK, Lewis, P & Thornhill, A. 2019. *Research methods for business students.* 8th ed. Boston: Pearson Education.

Schellnack-Kelly, I & Jiyane, V. 2017. Tackling environmental issues in the digital age through oral histories and oral traditions from the iSimangaliso Wetland. *Historia* 62(2): 112-129. https://doi.org/10.17159/2309-8392/2017/v62n2a6

Schellnack-Kelly, I. 2017. The role of storytelling in preserving Africa's spirit by conserving the continent's fauna and flora. *Mousaion* 35(2): 17-27. https://doi.org/10.25159/2520-5293/1544

Schellnack-Kelly, I. 2022. Research ethics to consider when collecting oral histories in wilderness areas such as the Kruger National Park. *HTS Teologiese Studies/Theological Studies* 78(3): a7467. https://doi.org/10.4102/hts.v78i3.7467

Schellnack-Kelly, IS. 2020. Decolonising the archives: languages as enablers and barriers to accessing public archives in South Africa. *Archives and Manuscripts* 48(3): 291-299. https://doi.org/10.1080/01576895.2020.1815064

Shibambu, A & Ngoepe, M. 2020. When rain clouds gather: Digital curation of South African public records in the cloud. *South African Journal of Information Management* 22(1): 1-9. https://dx.doi.org/10.4102/sajim.v22i1.1205

Shibambu, A. 2023. Strategies for developing digital infrastructure to rescue archives from stagnant archival holdings in the Government of South Africa. *Mousaion: South African Journal of Information Studies*. https://doi.org/10.25159/2663-659X/11661

Shukla, S. 2020. *Research methodology and statistics*. Ahmedabad: Rishit Publications.

Sigauke, DT. 2017. Digitisation technologies for Newspaper Archives in Zimbabwe: The ICT requirements for digitising a selected Bulawayo Newspaper Publication at the National Archives of Zimbabwe. IST-Africa 2017 *Conference Proceedings, International Information Management Corporation (IIMC).* https://doi.org/10.23919/ISTAFRICA.2017.8102315

Silva, AMDD. 2017. Folksonomies in archives: controlled collaboration for specific documents. *Ariadne 77*. Available at: https://hdl.handle.net/10316/40817 (Accessed 29 April 2024).

Silverman, D. 2020. *Interpreting qualitative data*. Los Angeles: Sage.

Silverman, D. 2022. *Doing qualitative research*. London: Sage.

Simons, GF. 2014. The role of metadata in the infrastructure for archival interoperation. *Language and Linguistics Compass* 8/11: 486–494. https://doi.org/10.1111/lnc3.12111.

Singleton RA & Straits, BC. 2018. *Approaches to social research.* 6th ed. London: Oxford University Press.

Skinner, J. 2014. Metadata in archival and cultural heritage settings: A review of the literature. *Journal of Library Metadata* 14(1): 52-68. https://doi.org/10.1080/19386389.2014.891892

Snyder, H. 2019. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research* 104: 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039.

Sobczak, A. 2013. *Traditional vs. Virtual Archives-The Evolving Digital Identity of Archives in Germany*. Poland: ProQuest.

Sonday, A, Ramugondo, E & Kathard, H. 2020. Case study and narrative inquiry as merged methodologies: A critical narrative perspective. *International Journal of Qualitative Methods* 19. https://doi.org/10.1177/1609406920937880

South African History Online (SAHO). 2019. https://www.sahistory.org.za/ (Accessed 20 October 2021)

Stangor, C. 2011. *Research methods for the behavioral sciences*. 4th ed. Belmont: Wadsworth Cengage Learning.

Stephanie, F & Ou, C. 2022. Revisiting oral history cataloging policies for multi-format discovery in MARC: A case study. *Technical Services Quarterly* 39(3): 225-240. https://doi.org/10.1080/07317131.2022.2082655

Strydom, A & Bezuidenhout, RM. 2014. Qualitative data collection. In: F. du PlooyCilliers, C. Davis, R. Bezuidenhout, eds. 2014. Research Matters. Cape Town: Juta, pp. 173-195.

Styblińska, M. 2006. LongTerm preservation of digital assets–Some specific aspects. In *Proceedings of the International Multiconference* 1896: 7094. Available at: Long-Term Preservation of Digital Assets – Some Specific Aspects (psu.edu) (Accessed 16 December 2021).

Sulehri, IG & Warraich, NF. 2020. Mapping the Metadata challenges in Libraries: A systematic review. Available at: https://digitalcommons.unl.edu/libphilprac/4679/ (Accessed 20 January 2022)

Surmiak, A. 2018. Confidentiality in qualitative research involving vulnerable participants: Researchers' perspectives. *Forum: Qualitative Social Research* 19(3): 12. https://doi.org/10.17169/fqs-19.3.3099

Swanson, R. 2015. Adapting the Brian deer classification system for Aanischaaukamikw Cree Cultural Institute. *Cataloging & Classification Quarterly* 53: 568-579. https://doi.org/10.1080/01639374.2015.1009669

Tani, A, Candela, L &d Castelli, D. 2013. Dealing with metadata quality: The legacy of digital library efforts. *Information Processing & Management* 49(6): 1194-1205. https://doi.org/10.1016/j.ipm.2013.05.003

Taniguchi, S. 2023. Data provenance and administrative information in library linked data: Reviewing RDA in RDF, BIBFRAME, and Wikidata. *Cataloging & Classification Quarterly*. https://doi.org/10.1080/01639374.2023.2178048

Tebeau, M. 2013. Listening to the city: Oral history and place in the digital era. *The Oral History Review 40*(1): 25-35. https://doi.org/10.1093/ohr/oht037

Teddlie, C & Tashakkori, A. 2009. *The foundations of mixed methods research: Integrating quantitative and qualitative techniques in the social and behavioral sciences*. Thousand Oaks: Sage.

Tembe, M & Nsibirwa, Z. 2022. Archiving the voices of the once voiceless: strategies for digital preservation of oral history at the KwaZulu-Natal Provincial Archives. *ESARBICA Journal* 41: 140-158.

Tembe, MA. 2019. *Preservation of, and access to oral history records at Pietermaritzburg Archives Repository*. Master of Information Studies, University of KwaZulu-Natal, Pietermaritzburg, South Africa.

Terre Blanche, M, Durrheim, K & Kelly, K.(eds). 2006. *Research in practice: Applied methods for the social sciences*. Cape Town: Juta and Company.

The Legal Deposit Act No. 54 . 1997. https://www.gov.za/documents/acts/legal-deposit-act-54-1997-14-nov-1997 ((Accessed 20 November 2021)

The Society of American Archivists. 2023. Dictionary of Archives Terminology. Last updated 2023. Available at: https://dictionary.archives.org/index.html (Accessed 20 January 2024)

Thomas, G. 2016. *How to do your case study*. Los Angeles: Sage.

Thomson, A. 2011. Memory and remembering in oral history. *The Oxford handbook of oral history*, pp.77-95. Ritchie, D.A. ed., 2012. *The Oxford handbook of oral history*. Oxford University Press.

Thor Tureby, M & Wagrell, K., 2022. Crisis documentation and oral history: Problematizing collecting and preserving practices in a digital world. *The Oral History Review 49*(2): 346-376. https://doi.org/10.1080/00940798.2022.2101933

Tinker, A. 2014. Creating order from chaos: a classificatory approach to retrieval in a local history photographic archive. *Journal of Writing in Creative Practice* 7(3): 427-445. https://doi.org/10.1386/jwcp.7.3.427_1

Tracy, SJ. 2019. *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. New Jersey: John Wiley & Sons.

Turhan, NS. 2019. Qualitative research designs: which one is the best for your research? *European Journal of Special Education Research* 4(2):124-136. http://dx.doi.org/10.46827/ejse.v0i0.2448

Turner, K. 2017. Creating history: a case study in making oral histories more accessible in the digital age. *Digital Library Perspectives* 33(1): 48-62. http://dx.doi.org/10.1108/DLP-06-2016-0016

Vaismoradi, M & Snelgrove, S. 2019. Theme in qualitative content analysis and thematic analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 20(3) http://dx.doi.org/10.17169/fqs-20.3.3376.

Van der Wal, E & Lambrechts, L. 2018. *Landscapes of loss: Curatorial mapping and the use of archival sound recordings. South African Music Studies* 36: 362-388.

Van der Waldt, G. 2020. Constructing conceptual frameworks in social science research. *The Journal for Transdisciplinary Research in Southern Africa* 16(1), a758. https://doi.org/10.4102/td.v16i1.758

Vasileiou, K, Barnett, J, Thorpe, S & Young, T. 2018. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC medical research methodology* 18: 1-18. https://doi.org/10.1186/s12874-018-0594-7

Venson, SL, Ngoepe, M & Ngulube, P. 2014. The role of public archives in national development in selected countries in the East and Southern Africa Regional Branch of the International Council on Archives region. *Innovation* 48: 43-64.

Vos, JN. 2013. *The development of a shared metadata standard for use in oral history collections.* Master of Science in Information/Library Science, University of North Carolina at Chapel Hill, North Carolina.

Wagner, C, Kawulich, B & Garner, M. 2012. Doing social research: A Global context. London: McGraw-Hill

Walker, I & Halvey, M. 2017.On designing an oral history search system. *Journal of Documentation* 73(6): 1281-1298. https://doi.org/10.1108/JD-10-2016-0121

Walker, Iain and Halvey, Martin (2017) On designing an oral history search system. Journal of Documentation, 73 (5). pp. 1-16. ISSN 0022- 0418 (In Press) , http://dx.doi.org/10.1108/JD-10-2016-0121

Walter, M. (ed). 2019. *Social research methods*. 4th ed. Australia: Oxford University Press.

Wamukoya, J. 2015. Reflections on African archives: Their role towards meeting the needs of society in the 21st century. *Journal of the South African Society of Archivists* 48: 14-20.

Weig, EC & Slone, M. 2018. SPOKEdb: open-source information management system for oral history. *Digital Library Perspectives* 34 (2): 101-116. http://dx.doi.org/10.1108/DLP-03-2017-0012

Wellington, J. 2015. *Educational research.* 2nd ed. London: Bloomsbury Academic.

Wiid, J & Diggines, C. 2013. *Marketing research*. London: Juta.

Wiles, R. 2013. *What are qualitative research ethics?* London: Bloomsbury.

Williams, J. 2016. The African Photographic Archive: Research and curatorial strategies. *African Arts* 49 (1): 95–96. https://doi.org/10.1162/AFAR_r_00279

Wood, S, Carbone, K, Cifor, M, Gilliland, A & Punzalan, R**.** 2014. Mobilizing records: re-framing archival description to support human rights. *Archival Science* 14: 397-419. https://doi.org/10.1007/s10502-014-9233-1

Woods, K, Chassanoff, A & Lee, CA. 2013. Managing and transforming digital forensic metadata for digital collections. *Proceedings of the Tenth International Conference on Digital Preservation (iPRES), Lisbon, Portugal,* September 2-6, 2013

Wynne, SC. 2009. Cataloging oral histories: creating MARC records for individual oral history interviews. *Cataloging & Classification Quarterly* 47(6): 561-582.

Xie, I & Matusiak, KK. 2016. Metadata. In Xie, I. & Matusiak, K. (eds), *Discover digital libraries: Theory and practice.* Elsevier, 129-170.

Yang, J, Chen, W, Feng, L, Yan, X, Zheng, H & Zhang, W. 2020. Webley supervised image classification with metadata: Automatic noisy label correction via visual-semantic graph. In *Proceedings of the 28th ACM International Conference on Multimedia* (MM '20), October 12-16.

Yang, LE. 2016. Metadata effectiveness in internet discovery: An Analysis of digital collection metadata elements and internet search engine keywords. *College & Research Libraries* 77 (1): 7-19. https://doi.org/10.5860/crl.77.1.7.

Yeo, G. 2012. The conceptual fonds and the physical collection. *Archivaria*:43-80. https://archivaria.ca/index.php/archivaria/article/view/13384.

Yin, RK. 2014. *Case study research design and methods.* 5th ed. Thousand Oaks, CA: Sage.

Yin, RK. 2016. *Qualitative research from start to finish.* 2nd ed. New York: The Guilford Press.

Zahle, J. 2021. Interpretivism and qualitative research. In *Stephen Turner and the Philosophy of the Social* (pp. 202-220). https://doi.org/10.1163/9789004449602_012

Zeng, ML & Qin, J. 2020. *Metadata*. American Library Association.

Zhang, A & Gourley, D. 2009. *Creating digital collections: a practical guide*. Oxford: Chandos.

Zierau, E & Peyrard, S. 2016. Digital preservation metadata in a metadata ecosystem. In Dappert, A., Guenther, R. & Peyrard, S. (eds.), *Digital Preservation Metadata for Practitioners.* Springer, Cham. https://doi.org/10.1007/978-3-319-43763-7_.

Zimu-Biyela, AN. 2022. Taking stock of oral history archives in a village in KwaZulu-Natal province, South Africa: Are preservation and publishing feasible? *HTS Teologiese Studies/Theological Studies* 78(3): a7438. https://doi.org/10.4102/hts.v78i3.7438

Zungu, M, Manqele, N, De Vries, C, Molefe, T & Hadebe, M. 2014. HERstory: Writing Women in South African History. *Agenda,* 28(1): 7-17. https://doi.org/10.1080/10130950.2014.871459

APPENDIX A: PARTICIPATION INFORMATION SHEET

25 May 2022

Title: Metadata schema of oral history content in South Africa.

Dear Prospective Participant

My name is Matlala Rachel Mahlatji, and I am doing research with Prof. Ngoepe, a Director for the School of Arts towards a PhD in Information Science at the University of South Africa. You are invited to participate in a study entitled metadata classification schema of oral history content in South Africa.

WHAT IS THE PURPOSE OF THE STUDY?

The purpose of this study is to explore the metadata classification schema of oral history content in South Africa, in order to propose a framework to ensure its access and retrievability.

WHY AM I INVITED TO PARTICIPATE?

As a staff member responsible for descriptive content for the oral history collections, you have been selected purposely based on these characteristics. The aim of this research is to focus on oral history content metadata and access; you are best positioned to share experience and knowledge to achieve the main purpose of research. I obtained your information from your Department. The approximate number of participants for the semi-structured interviews is ten, and the interviews will take place between the researcher and the staff.

WHAT IS THE NATURE OF MY PARTICIPATION IN THIS STUDY?

The study involves recording *semi-structured interviews* with the staff responsible for metadata creation of oral history content to understand the different ways in which archivists, librarians, and oral historians provide descriptive content for the oral history collections for which they are responsible. Your participation in this study will be to answer questions relating to metadata standards, metadata schema, access ad retrievability, technologies used, and issues relating to metadata classification of oral history content. This study aims to contribute to the proposal of a metadata framework to be adopted in the description of oral history collections to maximise the value and benefits of the collections for librarians, archivists, and citizens who use these collections. The duration of each interview will be approximately 60 minutes.

CAN I WITHDRAW FROM THIS STUDY EVEN AFTER HAVING AGREED TO PARTICIPATE?

Participating in this study is voluntary, and you are under no obligation to consent to participation. If you decide to participate, this information sheet will be given to keep and you will be asked to sign a written consent form. You are free to withdraw from the interviews at any time and without giving a reason. You may, however, not be able to withdraw any recording once data analysis has commenced.

WHAT ARE THE POTENTIAL BENEFITS OF PARTICIPANT IN THIS STUDY?

The study is expected to contribute to the already existing body of knowledge in LIS in the following ways.

- The study is likely to form the basis for further research on the metadata classification schema of oral history content nationally and internationally.
- The findings of this study will reveal the different ways in which archivists, librarians, and oral historians provide descriptions and metadata schemas, and which fields or elements each collection uses, and the ways in which the various fields or elements are similar and different across collections.
- The findings will also assist in establishing useful ways to promote access to and increase the use of oral history content and related materials in general.
- This study is expected to lay a foundation for a metadata framework to be adopted in the description of oral history collections nationally and more broadly.

ARE THERE NEGATIVE CONSEQUENCES FOR ME IF I PARTICIPATE IN THE RESEARCH PROJECT?

I do not foresee that you will experience negative consequences by participating in this study. The researcher undertakes to keep any information provided herein confidential, not to let it out of our possession, and to report on the findings from the perspective of the participating group and not from the perspective of an individual.

WILL THE INFORMATION THAT I CONVEY TO THE RESEARCHER AND MY IDENTITY BE KEPT CONFIDENTIAL?

This study will ensure that no one will be identified afterward. This will be done by recording the data in such a way that it cannot be linked to individual identification of the participants.

This means that the research will not record the names of the participant at any stage during the interview process. In cases where the participant's identity will be known to the researcher, the data will be de-identified and the identity is kept confidential (for example, interviews, where the participant identities are known to the researcher).

Additionally, access to data will be restricted as such only the research team (the supervisor, researcher, and research assistants), will not be restricted to access the raw data. The data collected in this research may be used for purposes such as research reports, journal articles, and conference proceedings. A study report will be submitted for publication, but individual participants will not be identifiable in such a report.

HOW WILL THE RESEARCHER(S) PROTECT THE SECURITY OF DATA?

Electronic copies of the transcripts will be encrypted and stored by the researcher for a period of five years on a password-protected computer. Future use of stored data will be subject to further research ethics review and approval, if applicable. After the five-year period, electronic copies will be permanently deleted from the hard drive of the computer. At the end of the research, on submission of the research report to the University, all hard copies will be shredded, and the audio-recorded information will be deleted.

WILL I RECEIVE PAYMENT OR ANY INCENTIVES FOR PARTICIPATING IN THIS STUDY?

Participants will not be reimbursed or receive any incentives for participating in the study.

HAS THE STUDY RECEIVED ETHICS APPROVAL?

This study has received written approval from the Research Ethics Review Committee of the College of Human Sciences. A copy of the approval letter can be obtained from the researcher if you wish.

HOW WILL I BE INFORMED OF THE FINDINGS/RESULTS OF THE RESEARCH?

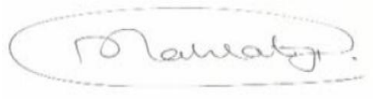
If you would like to be informed about the final findings of the investigation, please contact Matlala Rachel Mahlatji on 067 144 8624 or mahlamr@unisa.ac.za . The findings are available for five years.

If you need further information or want to contact the researcher about any aspect of this study, please contact Matlala Rachel Mahlatji on 067 144 8624 or mahlamr@unisa.ac.za.

If you have concerns about the way in which the research has been conducted, you can contact Prof. M Ngoepe 012 429 6360 or ngoepms@unisa.ac.za

Thank you for taking the time to read this information sheet and participating in this study.

Thank you.



Matlala Rachel Mahlatji

APPENDIX B: INFORMED CONSENT

CONSENT TO PARTICIPATE IN THIS STUDY

I, _____ (participant name), confirm that the person asking my consent to participate in this research has told me about the nature, procedure, potential benefits, and anticipated inconvenience of participation.

I have read (or had explained to me) and understood the study as explained in the information sheet.

I have had sufficient opportunities to ask questions and am prepared to participate in the study.

I understand that my participation is voluntary and that I am free to withdraw at any time without penalty (if applicable).

I am aware that the findings of this study will be processed into a research report, journal publications, and/or conference proceedings, but that my participation will be kept confidential unless otherwise specified.

I agree to recording the interview as a data collection method.

I have received a signed copy of the informed consent agreement.

Participant Name & Suriname…… (please print)

Participant Signature……. Date……

Researcher's Name & Surname…… (please print)

Researcher's signature……. Date.....

APPENDIX C: ETHICAL CLEARANCE LETTER

UNISA

COLLEGE OF HUMAN SCIENCES RESEARCH ETHICS REVIEW COMMITTEE

25 July 2025

NHREC Registration # :
Rec-240816-052
CREC Reference # :
30183162_CREC_CHS_2022

Dear Ms Matiala Rachel Mahlatji

Decision:
Ethics Approval from 25 July 2022 to 25 July 2027

Researcher(s): Name: Ms MR Mahlatji
Contact details: 30183162@mylife.unisa.ac.za
Supervisor(s): Name: Prof M Ngoepe
Contact details: 012 429 6360

Title: Metadata classification schema of oral history content in South Africa
Degree Purpose: PhD

Thank you for the application for research ethics clearance by the Unisa College of Human Science Ethics Committee. Ethics approval is granted for five years.

The *low risk application* was reviewed by College of Human Sciences Research Ethics Committee, in compliance with the Unisa Policy on Research Ethics and the Standard Operating Procedure on Research Ethics Risk Assessment.

The proposed research may now commence with the provisions that:

1. The researcher(s) will ensure that the research project adheres to the values and principles expressed in the UNISA Policy on Research Ethics.
2. Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study should be communicated in writing to the College Ethics Review Committee.
3. The researcher(s) will conduct the study according to the methods and procedures set out in the approved application.
4. Any changes that can affect the study-related risks for the research participants, particularly in terms of assurances made with regards to the protection of participants' privacy and the confidentiality of the data, should be reported to the Committee in writing, accompanied by a progress report.

University of South Africa
Preller Street, Muckleneuk Ridge, City of Tshwane
PO Box 392 UNISA 0003 South Africa
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150
www.unisa.ac.za

5. The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study. Adherence to the following South African legislation is important, if applicable: Protection of Personal Information Act, no 4 of 2013; Children's act no 38 of 2005 and the National Health Act, no 61 of 2003.
6. Only de-identified research data may be used for secondary research purposes in future on condition that the research objectives are similar to those of the original research. Secondary use of identifiable human research data require additional ethics clearance.
7. No fieldwork activities may continue after the expiry date (26 July 2027). Submission of a completed research ethics progress report will constitute an application for renewal of Ethics Research Committee approval.

Note:

*The reference number **30183162_CREC_CHS_2022** should be clearly indicated on all forms of communication with the intended research participants, as well as with the Committee.*

Yours sincerely,

Signature:

Prof. KB Khan
CHS Research Ethics Committee Chairperson
Email: khankb@unisa.ac.za
Tel: (012) 429 8210

Signature: PP

Prof K. Masemola
Executive Dean: CHS
E-mail: masemk@unisa.ac.za
Tel: (012) 429 2298

University of South Africa
Preller Street, Muckleneuk Ridge, City of Tshwane
PO Box 392 UNISA 0003 South Africa
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150
www.unisa.ac.za

APPENDIX D: DATA COLLECTION TOOLS

Annexure 1: Document Review Checklist

Document Title:			
Document reviewer (s):			
	Yes,	**No**	**Comments**
Guidelines for the creation of metadata for oral history contents			
Guidelines/ policies for metadata schema application			
Metadata schema/standards used for oral history content			
Metadata records: fields used / not used for oral history content			
Policy/guidelines for access to oral history at NFVSA			
Oral history content materials			

Annexure 2: Interview Guide for NFVSA Staff

'METADATA ***SCHEMA OF ORAL HISTORY CONTENT IN SOUTH AFRICA'***

Answer these interview questions as instructed below.

1. Please respond to the only questions asked.

2. Answer also follow-up questions.

INTERVIEW DETAILS

Interviewer:

Date:

Time:

Place:

Mode of interview:

Part A: Background Information

1. What is your job title or position at NFVSA?

2. What is your role in NFVSA?

3. How long have you been working at NFVSA?

4 What is your highest academic qualification?

PART B: Access to oral history content

5 In what format/s is oral history content at NFVSA?

6 How are oral history contents stored at NFVSA?

7 How many copies of oral history content material do you keep? (Duplicate and/or backups)

8 How are oral history contents made accessible at NFVSA?

8.1 Who has access to the oral history contents at NFVSA?

8.2 What are the search strategies are used to access oral history content in NFVSA?

9 How do you find metadata to enhance accessibility and discoverability of oral history content?

10 What are the problems associated with access and retrieval of oral history content?

Part C: Metadata standards/ schema

11 What metadata standards are used at NFVSA to classify oral history content?

12 Why are these standards used to classify oral history content at NFVSA?

13 How are metadata standards applied for the the classification of oral history content in NFVSA?

14 For what purpose do you use the current metadata? (Description, administration, and/or preservation)

15 Briefly explain the metadata creation process for oral history content at NFVSA.

15.1 What contextual information is collected on oral history content in the creation of metadata?

15.2 Describe the metadata fields that are used for oral history content at NFVSA.

15.3 Are there any metadata elements that are not applicable/used at NFVSA?

Briefly explain those not used and those missing from the current metadata schema.

16 What do you think can improve the gaps on the metadata standard/s used to classify oral history content?

Part D: Technologies used for oral history content

17 Identify the technological tools used for oral history content at NFVSA.

18 Explain which activities of oral history content metadata are done through technologies.

18.1 Why are those activities managed through technologies in NFVSA?

18.2 How are technologies used to access oral history content?

18.3 How are technologies used to create metadata for oral history content?

19 Why are these technologies important for the creation of metadata of oral history content?

PART E:

20 How often do you monitor and update changes to metadata?

21 What are the benefits of the use of the current metadata schema? Give a brief explanation.

22 What are the challenges of the current metadata schema for oral history content?

23 What are the limitations in the application of the current metadata schema for oral history. content at NFVSA?

24 Do you think the metadata schema used at NFVSA effectively assists with access to oral content?

25 Does the metadata schema application align to the policies at NFVSA?

25.1 How are these policies applicable in metadata at NFVSA? Give a brief explanation focusing on metadata creation, access and retrieval, choice of metadata schema

26 Would you recommend modifications/changes to the current metadata system? Elaborate by identifying areas for modification and what modifications should be done.

Society for American Archivists. https://dictionary.archivists.org/

National Archives of Records services of South Africa. https://www.gov.za/sites/default/files/gcis_document/201409/act43of1996.pdf

NARSSA 2016

https://www.nationalarchives.gov.za/ (access se 2024

The Legal Deposit Act 54 of 1997. 1997.

https://www.gov.za/documents/acts/legal-deposit-act-54-1997-14-nov-1997 (accessed

IASA https://www.iasa-web.org/