

cataloguing, instruction on archival description and preparation of a basic handbook. Recommendations for research and development were: research into how users obtain access to materials to provide a benchmark for the improvement of archival description; the definition of an archival information structure for effective administrative, physical and intellectual control; the evaluation and development of existing description practices; the examination of authority control to enhance user access; the development of guidelines for the use of controlled vocabularies; the integration of cataloguing rules for archival materials, special media materials, library materials and museum materials.

Recommendation 13 was particularly concerned with the definition of an archival information architecture for the "information management needs for the effective administrative, physical and intellectual control of archival materials".⁴² The variety of software applications in use at the time made the integration of information systems difficult. This applied to both internal systems in an institution, and the compatibility of systems between different institutions. This meant that, unless the same proprietary software was used, it was difficult to exchange information on holdings. In addition, where a finding aid and systems used for the management of holdings used differing technology, description had to be repeated across the two systems.

In 1989 a second edition of Hensen's *APPM* was endorsed by the SAA Council as a standard for archival description. Hensen's revision "shifted the emphasis from item level cataloguing to the description of collections (and) placed more emphasis on the descriptive areas relating to the content and analysis of the material and less on its physical aspects"⁴³. This provided a link between the broader standards for description and the standards for electronic coding. *APPM* made it possible for archivists to follow AACR2 rules and use its interpretation to create data values in MARC AMC. There were objections to the

⁴² Klaassen(ed.), 1989, p473

⁴³ Duff and Haworth, 1990-91, p26

role of *APPM* in the prescription of descriptive standards. One view was that *APPM* provides a "...de facto standard for archival descriptive cataloguing, more consistent with contemporary archival practice and compatible with library description"⁴⁴. The dissenting view, represented by Bearman, was that MARC AMC data content standards and *APPM* data value standards prescribed archival description principles. This, Bearman wrote, resulted from the constraints imposed by international description rules within which Hensen had to develop the manual, and because Hensen first set out to interpret the rules of *AACR2*.⁴⁵

The development of standards in the United States was heavily influenced by the technical requirements needed for cataloguing and data exchange. This can be seen by the retrospective approach taken. An exchange format such as MARC was developed, after which the limitations were identified, followed by adjustments to the format. The committees and working groups did not articulate a national view of how archival principles are applied to the development of standards. Description flows from arrangement, and other archival activities, which may or may not follow archival principles, but the main thrust of standards development focussed on linking the development of descriptive standards to the rules defined in technical standards such as *AACR2* and MARC AMC.

Canadian Standards Development

The Bureau of Canadian Archivists established a working group in 1983, to produce a set of proposals for adoption by the Canadian archival community, to develop standards and guidelines for the description of archival materials. Its scope included "the standards or guidelines developed by Canadian repositories to control descriptive practices and standards or guidelines developed outside the field of archival descriptive practice and outside Canada which were applicable to archival description".⁴⁶

⁴⁴ Fox, 1990, p22

⁴⁵ Bearman, 1992, p37-38

⁴⁶ Canadian Bureau of Archivists, 1985, p2-3

In 1985, *Toward Descriptive Standards: Report and Recommendations of the Canadian Working Group on Archival Descriptive Standards* was published and this laid the groundwork for the development of descriptive standards. Recommendations included the use of AACR2 rules for the formation of personal, geographic and corporate names and the development of authority files. In 1990, the first two chapters of *Rules for Archival Description (RAD)* were published. These covered general rules for description and rules for the description of multi-media fonds.

Towards Descriptive Standards defined archival description with reference to three of four major library descriptive functions: bibliographic description, choice of non-subject access points and subject indexing. It also aligned elements of description with AACR2, and *RAD* is based on the International Standard Bibliographic Descriptions (ISBDs), the basis of the American AACR2. No data structure standards or communications formats were endorsed. Canadian archival description was founded on the principle of *respect des fonds* and the application of the principle of provenance. Duff comments that "the development of descriptive standards and rules has refocused attention on the descriptive traditions and principles that govern archival description".⁴⁷

International Standards Development

The International Council on Archives provides the authoritative central point for the development of international archival standards. An Ad Hoc Commission for the Development of Descriptive Standards was established, but before standards were finalised, it was agreed that a statement on theoretical principles should be defined as a future foundation for greater acceptance of standards. In 1992, the *Statement of Principles Regarding Archival Description* was published after the review of the initial draft.⁴⁸

⁴⁷ Duff and Haworth, 1990-91, p29-33

⁴⁸ Duff and Haworth, 1990-91, p33

The General International Standard Archival Description (ISAD(G)) 2nd edition was published in 1999. It provides guidelines and general rules for archival description. These are described as part of a process that will "ensure the creation of consistent, appropriate, and self-explanatory descriptions; facilitate the retrieval and exchange of information about archival material; enable the sharing of authority data; and make possible the integration of descriptions from different locations into a unified information system".⁴⁹

The development of an international standard involves consultation with archivists worldwide. The 2nd edition was released after a five-year revision cycle that involved redrafting the *ISAD(G)* from a working document that included submissions for changes and the circulation of the draft for further review. The University College of London's research into the *ISAD(G)* was reported in *The Application of the ISAD(G) to the Description of Data Sets*, in 2000. The research project evaluated the *ISAD(G)* as a framework for the archival description of datasets. Three main issues were identified: problems with defining and describing metadata, problems associated with the cataloguing of electronic records separately from paper records, and problems with cataloguing electronic data sets at item level. This last problem was seen primarily as a technical one and not an *ISAD(G)* responsibility.⁵⁰

The Second Edition of the *ISAD(G)* was published in 2000, and its usefulness will be better known once archivists have had the time to apply and analyse it. The publication of the *ISAD(G)* has initiated the development of international standards and this has significance for South African archivists who have yet to establish a national standard.

⁴⁹ *ISAD(G)*, 2000

⁵⁰ Shepherd and Smith, 2000, p55, 85-86

Other National Standards Development Initiatives

In 1986, the Archival Description Project, funded by the British Library Board and the Society of Archivists, published their *Manual of Archival Description (MAD)* which proposed a descriptive standard for creating finding aids. *MAD* rejected the use of AACR2. British archivist, Michael Cook wrote that "it had not been possible to produce a standard for archival description which is adapted from AACR2", although he acknowledged that a national standard that could deal co-operatively with AACR2 was needed to facilitate data exchange.⁵¹ In 1990, the second edition of the manual, *MAD2*, was published. Walsh writes that in 1992 the intention was to develop a UKMARC AMC format. This would necessitate considerable effort in bringing existing online databases in line with the MARC standards and Walsh saw this as inhibiting the adoption of MARC.⁵²

Sweden's adoption of standards was preceded by the use of the draft version of the *ISAD(G)* to analyse Swedish application of archival principles and standards. The application of the principle of *respect des fonds* was found to be in accordance with the Swedish National Registry of Private Archives' arrangement and description rules. The concept of separate authority records was new to Swedish practices. The American MARC AMC format for the exchange of archival data was adopted.⁵³

Australia's use of the Australian Series System has been the subject of ongoing debate, as it dispenses with the traditional application of the principle of *respect des fonds*. Australian practices have moved rapidly into electronic recordkeeping. A project at Monash University, led by Sue McKemmish, was formed to establish a standard set of recordkeeping metadata standards. Adrian Cunningham describes the Australian view that "the objectives traditionally served by 'archival description' are but a subset of the broader set of objectives

51 Duff and Haworth, 1990-91, p27, cite Michael Cook, *The Management of Information from Archives*, (Brookfield, VT, 1986), p103

52 Walsh, 1994

53 Ottoson and Kristiansson, 1998

served by continuum-based recordkeeping metadata.” He describes the difference between the traditionalist and the Australian approaches to description as “static documentary or post-hoc bibliographic representations of records” compared to “the dynamic metadata regimes being explored in Australia.”⁵⁴

The differences in approaches to developing standards are marked. Canada and the United States made use of technology from the 1970s. Their administrative structures are notable for their stability and provide the infrastructure for the development and administration of standards development. Canada explicitly tied *respect des fonds* to its development approach while the United States did not. Australia has undergone a number of changes in government structures and, as a ‘young’ country, had to adapt rapidly. Its archival holdings are not founded on a long history of traditional archival practices in contrast to Britain, and it has implemented solutions that focus on current and future archival requirements. Sweden has also taken a practical approach, adopting technology solutions not previously used. Although Sweden’s archival holdings are some of the oldest, as a small country it deals with fewer large institutions, speeding up the process of consultation.

Encoded Archival Description (EAD) and Encoded Archival Context (EAC)

An important initiative in the development of an electronic encoded standard that is broadly acceptable to the international archiving community led to the establishment of the Berkeley Project, initiated by the University of California Library in 1993. The goal was to “investigate the desirability and feasibility of developing a non-proprietary encoding standard for machine-readable finding aids...”.⁵⁵ Important criteria in the development of the EAD included “1) the ability to present extensive and interrelated descriptive

⁵⁴ Cunningham, 1998

⁵⁵ The Library of Congress, 2001, *Encoded Archival Description (EAD) Official Web Site*

information found in archival finding aids, 2) ability to preserve the hierarchical relationships existing between levels of description, 3) ability to represent descriptive information that is inherited by one hierarchical level from another, 4) ability to move within a hierarchical information structure, and 5) support for element-specific indexing and retrieval.”⁵⁶

The development of the EAD Document Type Definition involved numerous specialists and organisations. In April 1995 the results were presented at a three-day Finding Aids Conference where the approach was accepted as offering a future standard. The Bentley Group, an association of the initial developers with experts in archival description and encoding techniques, was formed in 1995. In August 1995 the SAA Committee on Archival Information Exchange agreed to assist in involving interested archivists in the EAD DTD development. In 1996 an alpha or prototype version of the EAD DTD was released. The SAA Council submitted a formal request to the Library of Congress Network Development and MARC Standards Office to serve as the maintenance agency for the EAD DTD. This was followed in 1996 with a beta release of the EAD DTD. Workshops held to discuss this release generated the interest and involvement of a wide range of organisations. Under the sponsorship of the Research Libraries Group (RLG) and the SAA's continuing education programme more than four hundred archivists, librarians and systems administrators were introduced to the EAD. Workshops were held in the USA, the UK and Canada, providing an in-depth understanding of the EAD. At the end of 1997, the EAD Working Group reviewed fifty messages from beta testers throughout the world, including Britain, Sweden and Canada. In 1998, after revisions, Version 1.0 of the EAD DTD was released.

The history of the EAD DTD development involved a large number of interest groups, professional associations and national organisations. Extensive consultations, testing and research went into the project. The involvement of

⁵⁶ The Library of Congress, 2001, *Encoded Archival Description (EAD) Official Web Site*

archival institutions, particularly from the Northern Hemisphere is evidence of the need for a commonly accessible encoding standard. The adoption and implementation of EAD has been extensive in several of these countries. The EAD provides a common language in encoding archival collections. The Library of Congress acts as a central point for the maintenance of the EAD DTD. An online forum, the EAD listserv, allows EAD developers across the globe to communicate with each other, exchange information and debate the application of the EAD.

A further development in encoded archival description is the Encoded Archival Context Document Type Definition (EAC DTD). This extends the scope for contextual description by providing for information needed for authority records used by archivists to document and track originators and holders of archival records. The commonly used definition of archival context information is that it consists of "information describing the circumstances under which records (defined broadly here to include personal papers and records of organisations) have been created and used. This context includes the identification and characteristics of the persons, organisations, and families who have been the creators, users, or subjects of records, as well as the relationships amongst them".⁵⁷ EAC is intended to complement EAD and "support the descriptive needs of the archival community, specifically in the creation, maintenance, and publication of creator description...."⁵⁸ The revision of the *International Standard Archival Authority Record for Corporate Bodies, Persons and Families* or ISAAR (CPF) by the Committee for Descriptive Standards of the ICA is expected to accommodate the structure of EAC. The development phase of EAC is not yet complete. EAC falls outside the scope of this study, but reference is made to it where it has relevance to the application of archival theory in relation to descriptive practices.

⁵⁷ Ottoson, 2002

⁵⁸ Ottoson, 2002

Content, Context and Questions

Standards development does not start with the development of an encoded format such as EAD DTD. Reference has to be made to the application of archival principles, guidelines are needed for arrangement and description of archival material, and local guidelines need to reference national standards to format description into an encoded format. Different rules and standards may apply for varying description service requirements. Service requirements for description include user requirements for content and contextual information, access requirements, authority control requirements and exchange requirements.

Archival description in South Africa today has many similarities to that in North America during the 1970s. Local archivists and repositories have a parochial approach to their practices with no explicit agreement on standards outside individual institutions. The present study has been undertaken within this context. The scope is confined to the descriptive content and context necessary to fulfil user requirements to identify archival materials. Other service requirements such as access points, data exchange and control requirements are not analysed beyond the identification of relevant data components as they present in the research material. Successive chapters cover the following ground:

Archival Description in South Africa

The study will first establish the degree to which South African descriptive practices adhere to recognised descriptive standards. An electronic format places greater demands on archivists for standardisation. The only electronic standard used nationally is NAREM, which provides an Internet-based inventory of collections from a number of institutions with a prescribed data input form. As there is no formal national South African standard, the analysis of NAREM's display screen and descriptive content is done against the ICA's *ISAD(G)* as the first test for standardisation. EAD provides an encoding format for description input and not a content standard. While data input can be manipulated to fulfil formatting requirements, it is the content that conveys comprehensive meaning to

the user. The *ISAD(G)* offers an international content standard, which is largely accepted by the international archiving profession. The degree of compliance with *ISAD(G)* provides insight into common South African descriptive practices and is followed by an EAD markup in a later chapter.

Debates covering descriptive content have ranged from the minimalist approach to the anarchic.⁵⁹ My first question concentrates on the provision of data content and data elements within NAREM and their alignment with standards in the *ISAD(G)*. "Does NAREM's on-line screen content provide a basis for the provision of a user interface within the *ISAD(G)* standard?"

The EAD Standard and its Technical Application

The EAD DTD provides a format that promises greater flexibility than its electronic predecessors. The advances in technology should overcome previous limitations. The debates covering the application of archival principles create further demands on the versatility of encoding formats. The use of information technologies can provide the opportunity for the archivist to go beyond the traditional application of archival principles during the process of arrangement, description and markup. The application of the principle of provenance has proved to be contested. Its practical application, whether using the traditional or series systems, provide for differences in the accessibility and understandability of archival materials, while the maintenance of provenance within the hierarchy of description has been severely limited with previous encoding formats.

The second question is, "Does the EAD maintain the principle of provenance and can it accommodate the more dynamic contextual requirements needed in modern archival applications?"

⁵⁹ The Library of Congress, 2001, *Encoded Archival Description (EAD) Official Web Site*

A Case Study in EAD Markup

The relevance of the EAD to South African requirements, and the applicability of South African archival description to the EAD DTD is tested. The markup of the description of a South African collection to the EAD is done from two sources. The first, a detailed markup to item level, is from the paper-based inventory. The second tests the usability of the NAREM collection-level description for direct markup to the EAD. These two approaches identify the availability of appropriate data elements in the EAD, and the 'fit' between South African archival description and the EAD. The third question arises from this study, "Does South African archival description provide appropriate description for EAD markup?".

The Findings of the Research

The findings of the research results are drawn together. The implications of the study for descriptive theory will be drawn out. The conclusions reached and the results of the case study will be used to elaborate on the implications for South African application of the EAD. Finally, while the EAD provides a data structure standard, a data content standard must issue from local endeavours. A way forward will be suggested for both the technical markup requirements and for standardisation of descriptive techniques.

CHAPTER 2

An Overview of South African Descriptive Practices

The South African National Archives, NAREM and NAAIRS

In 1994 the election of South Africa's first non-racial government heralded far-reaching change in South African institutions. The restructuring of the State Archives' Service into the National Archives of South Africa brought changes in policy. With the National Archives Act of 1996, a mandate for a broader vision for the acquisition and the protection of materials was put into place. This, in part, reflects what Terry Cook describes as a move from a "juridical-administrative justification for archives grounded in concepts of state, to a socio-cultural justification for archives grounded in wider public policy and use".⁶⁰

The apartheid-based policies of the old State Archives translated into acquisition policies that focussed on the dominant white culture. Public records reflected the hegemony of the Nationalist Government. The collection policy for non-public records identified collections where "... veral klem gelê word op die dokumenteversamelings van persone wat 'n leidende rol op die gebied van die politiek, staatsadministrasie, plaaslikebestuur, regswese, ekonomie of krygswese gespeel het..."⁶¹. Organisations and clubs with similar qualifications were also considered, and active acquisition of papers from North America, the United Kingdom and Europe further entrenched the white Western culture.

The expression of South African culture and history in the State Archives was that of Afrikaner Nationalism with the addition of 'other-European'. The gap in the representation of the culture and history of the old political and social minorities is large. It is in the filling of this gap that the collections from private institutions are a valuable resource.⁶² The database for private collections, the

⁶⁰ Cook, 1997, p43-44

⁶¹ De Wet, 1984, p48-54

⁶² Kirkwood, 2002

National Register of Manuscripts or NAREM was established in 1978, a groundbreaking step in the provision of a central research resource. It provides a service to institutions and researchers with a central point of reference for private collections. NAREM and other databases held by National Archives are queried using the National Automated Archival Information Retrieval System or NAAIRS. In 2001, this facility was moved onto the Web, giving greater public access to these resources.⁶³

Outline of Analysis

The Canadian Working Group for Descriptive Standards, in their study of current practice, points out that “common sense dictates that future efforts to standardize descriptive practices will necessarily have to build on what now exists, an understanding of the configuration of current practice must precede any analysis of where and how standards may be developed.”⁶⁴

The accessibility of collection-level description from the NAREM database offers the opportunity to examine current description practices across a number of institutions. While National Archives defines the format of the description, its content depends on each institution's individual description practices. The study analyses this description in terms of content and context. In measuring the appropriateness of the NAREM display, the researcher's or user's viewpoint is referenced, within the context of archival theory and studies on description standards in the electronic environment. Three broad areas of analysis focus on metadata, element content and context. The *ISAD(G)* 2nd edition is referenced as an external standard against which NAREM description is compared.

NAREM was designed thirty years ago, without data exchange requirements in mind, technology was inflexible and prescriptive in its application and the development of description standards internationally was in its infancy.

⁶³ Kirkwood, 2002

⁶⁴ Bureau of Canadian Archivists, 1985, p35

This study has preceded recent changes in National Archives description practices. The National Archives has recently adapted the Canadian model of description, based on the *Rules for Application Description (RAD)*, for the description of public records. A workshop held by Terry Cook provided input into the development of revised description processes. A programme for the improvement of description of already described material and the publication of a standardised inventory will, in the future, enhance the resources available to the researcher.⁶⁵

Metadata and Meaning

The *ISAD(G)* describes its general rules for archival description as "part of a process that will:

- a. ensure the creation of consistent, appropriate, and self-explanatory descriptions;
- b. facilitate the retrieval and exchange of information about archival material;
- c. enable the sharing of authority data; and
- d. make possible the integration of descriptions from different locations into a unified information system."⁶⁶

While the publication of the *ISAD(G)* is relatively recent, countries involved in its development have themselves gone through a long process of description standardisation. The *ISAD(G)* rules are intended to enable data retrieval and exchange, and have been criticised for being prescriptive.⁶⁷ This criticism is not ignored, but adherence to the *ISAD(G)* requirements are rather taken as a starting point for a national database such as NAREM, whose possible future use could be extended to data exchange, as few of the contributing organisations have the systems to embark upon this.

⁶⁵ Calitz, 2002

⁶⁶ *ISAD(G)*, 2000, p7-9

⁶⁷ Hurley, 1993, p62-63

The following table details the results of an element mapping exercise between the NAREM elements and the essential *ISAD(G)* elements. The guidelines for NAREM data entry provide for additional comparison to *ISAD(G)* elements.

Table 1 - Element Map - NAREM to *ISAD(G)*

NAREM Element Name and Guideline	Equivalent <i>ISAD(G)</i> Element
Depot: The name of the institution where the collection is housed. This information only has to be entered on the first form. It will be duplicated for each entry when the data is captured.	Repository code (part of Reference code)
Type: The term 'MANUSCRIPT' is used throughout. This information also needs to be entered on the first form only.	Level of description If 'collection' is assumed
Reference: This is the number allocated to the collection or item when it is taken on stock. The number is important because it is used to trace the collection or item on the shelf. It is also used as a reference if an index to a particular guide is compiled.	Local reference code (part of Reference code)
Title or Description: A concise description of the collection. This is optional and can be ignored should it not serve a purpose for the particular institution. The description can indicate the subject or person referred to in the items, or the person who collected or donated them.	Title
Commencing Date/End Date: The period covered by the collection. If the day and month are provided the date is written as follows: 19381216 for 16 December 1938. Should only one date be stipulated, it is entered in both date paragraphs. If no period can be deduced, these paragraphs are left blank.	Dates consisting of Start creation date and End creation date
Remarks: According to requirement. It can contain, for example, the volume of the donation (number of folios, volumes, linear metres, etc.), whether an inventory exists, whether consultation is restricted by any special conditions, etc.	Immediate source of acquisition or transfer Physical extent Finding aids Conditions governing access

NAREM Element Name and Guideline	Equivalent ISAD(G) Element
Summary: Should the items deal with a particular person, a short biographical description can be provided. Thereafter follows a synoptic list of the contents of the donation (in order of series). ⁶⁸	Administrative/biographical history Scope and Content

While the guide provided by National Archives for the completion of a data entry form broadly defines the element content, a common understanding of the meaning and content of the elements facilitates consistency across all collection description. The essential elements for data exchange according to the *ISAD(G)* are reference code, title, creator, date(s), extent of the unit of description; and level of description.⁶⁹ An analysis of the NAREM elements that are candidates for mapping to the *ISAD(G)* details their appropriateness:

- The *ISAD(G)* reference code would use the NAREM Depot and Reference⁷⁰ but NAREM lacks a country identifier, which could be created by the system underlying NAREM. In addition, the ISO 3166 codes for the representation of names of countries is the standard required by *ISAD(G)*.⁷¹
- The guidelines for NAREM data entry provide for an optional Title as opposed to an essential entry as stipulated by *ISAD(G)*. The Title at the collection level also needs to include the name of the creator of the records.⁷² The creator is a separate element, and is used in the *ISAD(G)* Context Area.⁷³
- While the Depot names on NAREM are unique, the establishment of a national repository code standard would reinforce uniformity across all forms of collection-level description, and not only in NAREM.

68 National Archives of South Africa, undated guidelines, *The National Register of Manuscripts (NAREM)*

69 *ISAD(G)*, 2000, p7-35

70 The NAREM elements are referenced as proper nouns to avoid the use of inverted commas.

71 *ISAD(G)*, 2000, p13

72 *ISAD(G)*, 2000, p14

73 *ISAD(G)*, 2000, p18

- The Commencing Date/End Date provides for the "period covered by the collection" and for, both commencing and end date, or a commencing date. This is similar to the *ISAD(G)* rule but, on closer inspection, the content is subject to a broader interpretation.⁷⁴
- NAREM provides for the physical extent for the unit of description in the Remarks element. This is the equivalent of the *ISAD(G)* requirement⁷⁵ but as there is no separately defined element in NAREM, its inclusion depends on the information supplied by the institution concerned.
- The element Type appeared to indicate the level of description, where a uniform setting of "manuscript" would indicate a manuscript collection-level description. But the setting varies and in some instances the description content does not relate to collection-level description.⁷⁶

The NAREM elements do not qualify fully as essential *ISAD(G)* elements.

NAREM Content and Context Analysis

In the *Introduction* to the *ISAD(G)*, the standards are described as providing general guidance. Further, they are "to be used in conjunction with existing national standards or as the basis for the development of national standards".⁷⁷ National standards for content such as Hensen's *Archives, Personal Papers, and Manuscripts: A Cataloguing Manual for Archival Repositories, Historical Societies, and Manuscript Libraries (APPM)* used in the USA, Canada's *Rules for Archival Description (RAD)* and England's *Manual of Archival Description (MAD)* provide this guidance. With no national description standard in

⁷⁴ This is discussed in the analysis of content in Table 2

⁷⁵ *ISAD(G)*, 2000, p18

⁷⁶ The analysis of Table 3 provides an example of item-level description, which appears to be for a collection.

⁷⁷ *ISAD(G)*, 2000, p7

South Africa, common practices in description can contribute to the development of a national standard.

The application of a standard relies on consistent use of the data elements with appropriate content. A content analysis was done of the elements in the NAREM database. The viewpoint of the researcher was taken in looking for consistency in the use of elements and the usefulness and completeness of content. Three collection-level descriptions were extracted using 'African National Congress' as a search string. These examples from the University of Natal, the University of the Witwatersrand and the University of South Africa (UNISA) are compared in Table 2. Collection-level description was deliberately chosen as it fulfils the conventional and intended use of NAREM as a collection-level database. To highlight important features of the description, NAREM Remarks and Summary elements which allow a free-format entry, have been shortened.

Table 2 - Description Content Comparison⁷⁸

Field Name	Document 1	Document 2	Document 3
Depot:	Wits University Library	Alan Paton Centre and Archives, University of Natal, Pietermaritzburg	UNISA
Type	Manuscript	Manuscript	Accession
Reference	AD2186	PC31/1	AAS46
Title or Description	African National Congress (ANC)	African National Congress collection.	JS Moroka Accession
Starting and Ending Dates	1928 1990	1994 1994	1941 1974
Remarks	Records; 12 boxes; inventory. Forms part of the archives of the South African Institute of Race Relations.	12 file groups. Access unrestricted.	Donor: Dr. JS Moroka; 0,10m. Inventory
Summary	+ Constitution, minutes, reports, resolutions, Treason Trial legal records, ...	The ANC was inaugurated in 1912... In the 1950s ... + PC31/1-11 Policy	The donor, Dr. James Sebe Moroka (1891-) was born in Thaba'Nchu ... President-General of the ANC 1949-1952.

⁷⁸ The examples in the table can be viewed in Appendix A. The term 'document' is used for the displayed collection description as this is the convention for NAAIRS.

Field Name	Document 1	Document 2	Document 3
	+ Subjects are apartheid, African nationalism, passes, education, Freedom Charter and the activities of the ANC in exile. + There is information on Albert Luthuli ...	framework drafts, statements, manifestos, publications, pamphlets, voter education material, election posters. ... <i>(display truncated)</i>	+ Addresses, articles, circulars, correspondence, minutes of meetings, notes, ... + Professional papers, including Moroka Methodist Hospital, ... <i>(display truncated)</i>

An important means of providing context is the hierarchical structure of the collection arrangement and additional contextual information. These examples illustrate aspects of contextual description. In NAREM the Reference element is used by the institutions to identify their collections. Document 2 indicates this hierarchical arrangement as the Reference, PC31/1 implies that this description finds its context within a higher level of description, PC31. The Reference number in Summary goes down a level (PC31/1-1) for a lower level of description, reflecting existence of the file groups described in Remarks. Document 1's contextual Remarks, situates the collection within the Institute of Race Relations' archive. Document 3 positions the creator or collector of the records as a past president-general of the ANC.

The description content and sequencing of the description components in the Remarks and Summary elements differ across the collections.

- Physical extent, arrangement, access conditions, source of acquisition, type of documents, and the availability of an inventory are in some but not all the Documents.
- In Summary, Documents 2 and 3 provide a history or biography of the creator of the records followed by a description of the collection, by file group. The *ISAD(G)* Context Area accommodates these components

with comparable historical and biographical elements describing the creator of the records.⁷⁹

- The description of file groups or series within Document 2 follows the rule that description is defined from the general to the specific.⁸⁰
- Biographical, historical and series description components comply with NAREM's input guidelines for the Summary element.
- In contrast, Document 1's description reflects types of records, subjects and important people and organisations. These components are catered for in the *ISAD(G)* Content and Structure Area.

The two approaches to description result in content that fit into aspects of *ISAD(G)* requirements.

The researcher's understanding of the description is assisted by consistency in the use of elements. The description content of each element must reflect the rules defined for the element, which must be precise to avoid confusion. The use of Date is an example of an element that is not clearly understandable by the researcher. In Document 2, Date is used differently to the other examples and could refer to the date of the creation of the description, the acquisition of the collection or the markup to the NAREM database. Documents 1 and 3 Date/s probably reflect the document dates, but the researcher needs clarity as to the type of date displayed.

An example of a permissive and potentially confusing data definition is that of 'Type' which could either relate to the physical medium such as photograph, paper document, or electronic record, or the function such as letter, journal or financial report. The inter-changeable use of this type of element is confusing for the researcher and does not aid electronic access, where search values may be assumed to fall within either of these categories, with a loss of contextual

⁷⁹ *ISAD(G)*, 2000, p18

⁸⁰ *ISAD(G)*, 2000, p12

information. Ideally, if description is maintained for both definitions, two appropriately labelled elements should be clearly distinguishable.

Cross-Database Analysis

The intention of this part of the research was to analyse entries on the NAREM database, but this was extended to include description from other databases. When taking the researcher's view, it was not immediately apparent that a query could result in a cross-database selection, including records from other databases. Unless a selection is made on 'type' in the query facility, a number of databases are accessed including the National Registers of Manuscripts and Photographs (NAREM and NAREF) and the National Archives' cartographic and library material, microfilms and copies. This results in the researcher needing to navigate through different display formats, with differing combinations of elements, unaware that these entries are not from NAREM.

An understanding of the content of an entry is predicated on knowledge of the 'identifiers' or codes, Source, Type, Volume No, System, Reference and Part, and the rules governing their interrelationship. Initially, two basic types of entries were identified. The first was the non-public records, which are held at private institutions, and use the Type and Reference elements. This subset of elements provides a simpler display than that for public records. However, some material falls under the category "non-public" material, which is in the National Archives or provincial repositories and also forms part of the NAREM database. These materials fall outside the scope of 'official' records but relate to collections and other materials such as journals, pamphlets, etc. created in the 'non-public' sector. The researcher can narrow down the selection of Type to one of cartographic material, copies, library material, microfilm, non-public records, photographic and pictorial material. Public records, the second type, are described using a mix of all the identifier elements listed above. It was found after tracing through the meanings associated with the Identifier elements, that the non-public records held at the national or provincial repositories have a hybrid set

of element rules from both public and non-public records. The lack of immediately interpretable description that would locate a document in the correct database and domain (public, private, or publicly-held private material) hides pertinent contextual information. Extracts of a search for ANC-related documents are contained in the following table and illustrate specific aspects of this analysis. The shaded areas are elements that are not used in the relevant document.

Table 3 - Results of Cross Database Query⁶¹

Document	4	5	6	7	8	9	10	11	12
Element									
Depot					NALN	TAB	VAB		
Source						Biblioteek			
Type	Manuscript	Accession	Manuscript	Manuskrip	Manuskrip	Pamflet	Biblioteek	Manuscript	Manuskrip
Reference	AL2448	AA6268	329 9 (68)	M92136/93/ 731	M92136/93/ 732	P748	Tydskrif	AL2421	
Part						17			
Description	Portraits of ANC Leaders			Prakt met die ANC	Prakt met die ANC	African National Congress ...	Argentineus Dkt 1892	Natal Indian Congress	RGN Versameling
Starting			0000				0000		0000
Ending			0000				0000		0000
Remarks	8 paintings			Gedeeft van die Taurus- manuskrip versameling ...	Gedeeft van die Taurus- manuskrip versameling ...			...South African Tin Workers Union (AL2420),...	... Brief... nr 5838
Summary	... paintings ...	... leaflet	Newspaper articles						

The description includes codes and abbreviations that can be translated from lists on the National Archives web site, but there is no 'contextual' help on the document web page itself.⁶²

- Depot is the archives repository at which the archives concerned are housed
- Source provides the entities or sources according to the office that created or received the material

61 National Archives of South Africa Web Site, 2002
62 National Archives of South Africa Web Site, 2002

- Type is the physical type of the relevant archives.
- Volume no and reference enables the identification of the physical location of an item.
- Description is a synoptic description of the archival unit
- Starting and ending dates.
- Remarks providing explanatory details such as cross-references are given here.
- Summary is used only in data pertaining to the national registers of non-public records.

If these explanations are not accessed, and in this case, they were not accessed until it became apparent that the ascribed meaning differed from the actual use, the result is a misunderstanding of the content of the elements. The use of abbreviations instead of full names for Depot results in a loss of information to the researcher. TAB (meaning the National Archives Repository, Records of the former Transvaal Province and its predecessors as well as of magistrates and local authorities) and VAB (Free State Archives Repository) are public repositories while NALN (full name not found on the web site), is a private organisation. This information indicates where the records were created or used and informs the researcher of their physical location. This distinction is also important in the interpretation of other elements.

The Source, Type and Reference elements have rules that differ according to dependencies between elements based on changing content. Type does not refer to the documentary form or the medium of the materials as may be interpreted by a researcher. It can contain "manuscript", "manuskrip", "accession", "aanwinst", "private collection" and "private versameling", which are synonymous and refer to the NAREM database of private collections. The Type "pamflet" refers to "library material in the custody of national and provincial archives repositories".⁴⁹ The same values in different elements are also a source

⁴⁹ National Archives of South Africa Web Site, 2002

of confusion. The Source "biblioteek", refers to the library at the TAB repository (Document 9) according to the above rule while the Type "biblioteek" is not explained, but has to be assumed as the library at the VAB repository (Document 10).⁸⁴

The Canadian Working Group found that, "Archivists have not worked out generally accepted terminology to distinguish form as related to medium from form as related to function".⁸⁵ In these documents, medium and form are not distinguishable as defined elements but are rather scattered across Type, Summary, Remarks, and Reference. The Reference element that should provide a unique Identifier does not have a consistent rule, as is seen in Document 10 with a Reference of "Tydskrif". The value, "Pamflet" can be accessed in Type but not paintings, leaflets and newspaper articles as these are of the Type "Manuscript". These are found instead in the narrative of Summary. The term "Pamflet" if used as a Type has an extended rule that locates it as "library material in the custody of national and provincial archives repositories" and this excludes this use for pamphlets that are housed in private repositories.⁸⁶

A coherent view of related documents provides the contextual view necessary for the researcher. However, the first barrier to this is that of the use of regional languages, in this case, Afrikaans. This illustrates old practices that relate to the period where Afrikaans was the preferred language of government. Access to the on-line description, in a mainframe 'green screen' environment, was initially only available at regional offices where the staff could assist in interpretation. Later this was extended to some South African institutions. As Afrikaans was one of two official languages, these institutions had South African personnel familiar with Afrikaans. This information would not be accessible to non-Afrikaans-speaking South Africans. Further, the availability of the NAREM database to a broader international audience through the Internet creates an

⁸⁴ Further investigation using type library failed. The document was found on a search on all the types.

⁸⁵ Bureau of Canadian Archivists, 1985, p43

⁸⁶ National Archives of South Africa Web Site, 2002

imperative to provide translations and search facilities on the translated description.

The hierarchical structure of the levels of description for a collection creates further context, with the inheritance of description from the highest down to the lowest levels. As NAREM is not intended to provide more than a collection-level description, the attempts by institutions desiring display at a more detailed item level results in multiple entries for the same collection, in some cases. The same Description and Reference Numbers in Documents 7 and 8 point to consecutive items in the same collection. However, there is no mention of this relationship or the level of description in either document. This parallels a bibliographic form of encoding where items are discrete, needing no reference to other items for contextual completeness. Beaman, in examining the MARC AMC format, differentiates between bibliographic and archival methods of description. The librarian, he writes, "*transcribes* cataloging information from the *item in hand*", while for the archivist the items are described in terms of "*the activity out of which they arose and the actions that have been taken on them*". The MARC-AMC format was created to accommodate the requirements of a community that described cultural materials without considering their physical format as a primary feature.⁶⁷ Some of the Documents described at discrete item level do not even fulfil the criteria for bibliographical cataloguing and are not identifiable as discrete items. Document 12 has no unique Reference and when a search was done using the Title, five documents with the same Title were identified.

The relationship between Document 12 and these 'sibling' documents is not apparent as there is no reference to a 'parent', main heading or series header. There are also no cross-references from one to another. Some documents contain references in their narrative description. These are unusable in an automated environment as they are not discretely defined elements that can be programmatically accessed. The numbering of Reference on Document 7

⁶⁷ Beaman, 1989, p29-39

and 8 indicates the possibility of an hierarchical arrangement. A query string based on the description, "Praat met die ANC", resulted in five entries with reference number suffixes from 731 to 737. It is unclear whether these item-level descriptions are contextually related or if the reference numbers are a list with no significance attached to the prefix, other than physical location.

There is a lack of cross-referencing between related records so the form of arrangement is not transparent to the researcher. This situation necessitates the intervention of an archivist with 'inside' knowledge of the system, but it also needs the researcher to initiate an enquiry. A complex, coded form of description that does not reveal the hierarchical structure of arrangement creates the possibility that the 'gaps' will go unnoticed and material sources missed. The Canadian Working Group comments that in their investigations into Canadian practices " ... It is not certain whether conscious identification of the various levels is made during the construction of finding aids."⁸⁸

The date elements, which describe the period of the creation and use of the materials, offer further context. Document 10, which describes a journal, a library item, uses a discrete bibliographic type of entry. It does not make use of the Date element, using instead Description. This excludes this journal from a search based on Date.

Additional information that aids the researcher includes the availability of an inventory, contact and reference information for accessing the material and any restrictions placed on access. The name of the repository and contact information is available the list on the National Archives web site titled "Acronyms and names used to identify archival repositories (depots) and contact details."⁸⁹ The availability of an inventory is not explicit as some repositories may have an inventory but do not refer to this in the description, as there is no element or prescriptive rule on NAREM for this information. Access restrictions and

⁸⁸ Bureau of Canadian Archivists, 1985, p43
⁸⁹ National Archives of South Africa Web Site, 2002

conditions are rarely provided. The legislation affecting access, terms of access, period of access, and the date when the status of the item changes were not found on public records, some of which have dates as far back as the 1950s, and it is not clear if access restrictions remain in force.⁸⁰ Some description is so minimal that all that can be seen is that the material is about the ANC. The purpose of the finding aid appeared to be the aiding of user to physically access the materials, rather than providing enough information to decide first if the materials are relevant to the scope of research.

The linking of materials to provide context goes beyond the ability to view or reconstruct the arrangement. A collection has a further context in its cultural and historical connections. References in the material to significant events, organisations, people, etc can point to further context. NAREM is a resource rich in contextual relationships between collections. Research would be considerably enhanced if institutions provided cross-references to related entries. Document 11 illustrates this, with the inclusion of the collection identifier for the collection of the South African Tin Workers' Union, AL2420.⁸¹ This is an 'internal' reference to a NAREM collection. Linked access to the web page of the NAREM collections, institutions' repository web pages and other online archives and publications further expand the horizon of context. The boundaries of contextual information appear unlimited.

The context of creating the ANC collection was not isolated from the political and social upheavals of the apartheid era. Many records of the ANC were periodically confiscated and destroyed, and the archivist was a political activist, subjected to interrogation and harassment. The view that the archivist practices as a disinterested professional, resulting in description 'untainted' by the events surrounding the creation of the records depletes the archive of further context. The */SAD(G)* element, Archivist's Note provides an example of how this

⁸⁰ See Appendix A, Document 14 for an example.

⁸¹ Appendix A

information can be described. In the case of the ANC archival, a collection description is available on NAREM (Document 13).⁹² It provides additional context supporting the statement in the hardcopy inventory, "In 1979 it was transferred to the library by the Institute's archivist, the late Jeanette Curtis (Mrs. Marius Schoon, who was assassinated in Angola in 1984) to ensure its safe-keeping"⁹³. This enhances the contextual knowledge of the researcher, who understands then the political context within which the original archivist worked, how her perspective of the collection and its accompanying description was influenced by the prevailing circumstances.

Conclusion

The results of a survey described in "Retrieval, Accessibility, and Preservation as Reflected in the Questionnaire on Manuscript Collecting" published in 1984 identified that, of finding aids for public access to manuscripts, 65 holding organisations used catalogue cards, sixteen used a 'book' form and three used computerised resources. Of twenty replies, two-thirds used AACR2 rules for cataloguing. In making the contents of manuscripts available, thirty-five institutions used inventories and guides, while twenty-seven institutions replied that there were not the resources to compile these finding aids.⁹⁴ There is no quantitative data on the latest level of computerisation in the archiving sector, but NAREM remains the only Internet-based register offering access to South African collections and provides an invaluable resource for the study of description practices.

The collection-level description in NAREM generally offers the researcher sufficient content to understand the individual entry. The design of the layout and the use of free-format text areas for important description components results in a lack of consistency in providing all the information needed. If some information

⁹² Appendix A

⁹³ *Records of the African National Congress Collection, 1992*

⁹⁴ Cartwright, 1985, p26-32

is excluded, it is not evident whether this is because it is not applicable, an example being whether an inventory is available. This is complicated by the interchangeable use of terms, some which combine archival description components such as function and form. The broader context viewed through the hierarchy of arrangement is not readily identifiable because levels of description and method of arrangement are not explicit and there is no internal reference to these relationships. However, where the approach to description is a single entry per collection, there is no need for this view.

Across database searches, particularly including public records, are more complex. Codes and abbreviations are used and their metadata rules have interdependencies, which are not transparent to the researcher and deprive him or her of contextual information. Their translation is complex, and the internal rules are not always adhered to with resulting misinterpretation. The identification of levels of arrangement and related items is particularly difficult and sometimes impossible. The codes relate more to the physical location of the material rather than the form of arrangement. There is no room for more dynamic description such as a change in access conditions or the highlighting of additions and changes to the collection. The identification of public as opposed to private records and the database on which the item resides is not readily visible.

The conclusion reached by the Canadian Working Group in their study holds similarities to the findings of this study. The Canadian Working Group found "descriptive... practice is highly idiosyncratic, ... archivists capture virtually the same types of information about their holdings."⁹⁵ The similarities between the state of current South African archival description and that of North America in the 1980s are apparent, as descriptive practices were parochial, and while there were similarities in the kind of information provided, there was little standardisation as to the format and use of data elements. The Canadian Working Group identified problems in the consistent application of terminology,

⁹⁵ Bureau of Canadian Archivists, 1985, p53

the use of the finding aid, the inclusion of levels of arrangement and how different materials are included in the finding aid.

The common set of assumptions predating Canadian and American archivists' approaches to developing descriptive standards were: the principle of *respect des fonds*, archival description proceeds from description of 'collectivities' of archival material to lower levels. While description focuses on records and private papers, in addition, description of the creator of archival materials, as well as functions, activities and transactions from which the records emanated is necessary.⁹⁶

The maintenance of the principle of provenance relies on a system of description that reflects the hierarchical structure of the materials, providing links between higher and lower levels of description. These need to be readily accessible and viewable by the researcher, not relying on an unformatted display or extrapolating the possible order through the deconstruction of codes and searching for entries that may be related. While the individual collection descriptions are appropriately self-explanatory, some contextual issues are raised when looking outside the collection-level description. The ANC Collection, housed at the University of the Witwatersrand, is used as a case study. It was selected as it is accessible at Historical Papers in the William Cullen Library, University of the Witwatersrand. It is also a collection with a detailed hardcopy inventory and an entry on NAREM, so both the collection-level and detailed description is available for this research. In contrast, the official ANC Collection is not described on NAREM and is located at the University of Fort Hare. Context for the University of Witwatersrand ANC collection is provided as it states that it is part of the Institute of Race Relations archive, while the source of the University of Natal's collection is not provided. The Encoded Archival Context DTD and the *ISAAR(CPF)* provide the possibility of identifying the full range of collections. Until there is co-operation in making collection description centrally viewable, a

96 Duff and Haworth, 1997, p198